

2041
GCCAGGCGAAGAACTTCCCTGTCCTGTTTTCACAGACAGGTTCTCTGGAGAAAGGAGATGTTCTGCCCAAGAGGGGGCCGCACTCTGAGCAGAAAGCTGGAAACAAGTGCACACTCAAAACCCGCG
A G Q K A S L S V S Q T G S W R R G M S A Q G G A P S R Q K A G T S A L K T P G

2161
GAAACCGATGATGCCAAAGCTTCTGAGAAAGAAAGAGTCACTCTCTCAACAAGATCTCTTCAGATGCAAGAAAGACAGTGAAGGAAAAAGCCCCCTC
K T D D A K A S E K G K A P L K G S S L Q R S P S D A G K S S G D E G K K P P S

2281
AGGCATTGGACAGATCGACTGCCACAGCTCTTGGCTTTAAGAAACCAAGAGTGAAGGATCTCTCCATGATACACAGCAAGTGGAGCAACCATATAACAAGTGGCTCTGCAACAACCTGGG
G I G R S T A T S S F G F K K P S G V G S S A M I T S S G A T I T S G S A T L G

2401
TAAAAATTCCAAAATCTGCTGCCATTGGCGGGAAGTCAAAATGCAGGAGGAAAAACAGTTTGGACGGTTCACAGAAATGAGATGATGTTGGCTGATGTATTAGCTCAAAAGACTCAACCTTACA
K I P K S A A I G G K S N A G R K T S L D G S Q N Q D D V V L H V S S K T T L Q

2521
ATATGCGACAGTTCGCCCGCCCTTCAAAATCCAGACACAGTGGCACTTCTGCGCGAGAGAGCCACAGATCCAGTACAGAGATTTGATTCCAAAGCTCAGCAGCAAGTCTGCGGGCCAC
Y R S L P R P S K S S T S G I P G R G G H R S S T S S I D S N V S S K S A G A T

2641
CACCTTGAAACTGAGAGAACCAATTAATTGGGTGAGGCGGCTCGAGTCTGTCTACCGTTCACCAACAGACAGGAAAAAGAAAGATGACGACTTCAGATTCCAGAAAGTGTCTTT
T S K L R E P T K I G S G R S S P V T V N Q T D K E K E K V A V S D S E S V S L

2761
GTCAAGTTCGCCCAATTCAGGCCCCACACTCTGCGCAGCGCTGTGGTGCACAAGGTCTCAGGCAAGCAGATCCAGATTCAGATTTAGCTTACCCACATTTTGAAGGTGTGTTGGTGC
S G S P K S S P T S A S A C G A Q G L R Q P G S K Y P D I A S P T F R R L F G A

2881
CAAGCAGGTGGCAAAATCTGCTCTGACCTTAATCTAGAGGTGTGAATCTTCTTCAGTAAATGAGTCCCAAGCCCTTAATGATACCAATTTAGCGCGCAAGGCAAGTCTGAGTCAACGCTGTCCGG
K A G G K S A S A P N T E G V K S S S V M P S P S T T L A R Q G S L E S P S S G

3001
TACGGGAGCATGGGCACTGCTGGTGGGCTAAGCGGACAGACGCCCTCTCTTCAATTAACCTCAGACTTAATCTCAGATGTTATTAAGCTTAAAGTCACTCGTGGCTCCAGCCCAAGC
T G S M G S A G G L S G S S S P L F N K P S D L T T D V I S L S H S L A S S P A

3121
ATCGGTTCACTCTTTCACATCAGGTGGTGTCTGTGGGGTCCCAATTAATGAGCAGTTCCTCTGACAGGACGACGAAAGATTTCTCGGACGTCACAGTCCATGACTAGCCTCCACACAGAGCTCTGA
S V H S F T S G G L V W A A N M S S S S A G S K D T P S Y Q S M T S L H T S S E

3241
GTCCATTGACCTCCCTCCAGCCATTCAGTGGCTCTGTGTGAGCTGACCAAGACACTCAGAGGTTCAGAGGCTGCTCATGAGAAACGGGTAGTGTGAGTCTCTCTCAGAAAACAT
S I D L P L S H H G S L S G L T T G T H E V Q S L L M R T G S V R S T L S E S M

3361
GCAGCTTGACGAAATTCACATACCTTACCAAAAGGAGCTTAAGATTTTACCCCATTCATCTCGGACGAGCCCAACCAAGAGGGCAAGAGTGGTGTGCTCTCATTTCTGAGAGGCGCTTCAGGA
Q L D R N T L P K K G L R Y T P S S R Q A N Q E E G K E W L R S H S T G G L Q D

3481
CACTGGCAACCAAGTCACTCTGTGTTTCCCTTCTGCAATGTCATCTTTCAGAGCTGAGAAATTCACACTTTCTTAATCTTGAGAGCCCAACAATTTGTCTCAATTTTAACTTCCCGGGCC
T G N Q S P L V S P S A M S S S A A G K Y H F S N L V S P T N L S Q F N L P G P

3601
CAGCATGATGGCTCAACAGACATCCCAAGCCCAAGACTCTTCCTTGATCTCTATGATGATCCAGCTTTGTGGAGTGGCCACTTCTCTGAGGAAAGACCTGTGCATCAGTCAATTC
S M M R S N S I P A Q D S S F D L Y D D S Q L C G S A T S L E E R P R A I S H S

3721
GGGCTTATTCAGAGACAGCATGGAAGATTCATGGCTCTTCAATTATCACTGGTGTCCAGCACTTCTTCTTACTTACGCTGAAGAAAGGCTCATTCAGAGCAAAATCCATAACT
G S F R D S M E E V H G S S L S L V S S T S S L Y S T A E E K A H S E Q I H K L

3841
GGGAGAGAGAGTGTGCTGCATCAAGAAAAAGTGTGTCACCTTCACATCTCAGCTTTCAGCAAAATGCTCACTTGTAGCAGCTTTTGAAGAGCTTAAAGGAAATGATGATGCGCGGATTTGCA
R R E L V A S Q E K V A T L T S Q L S A N A H L V A A F E K S L G N M T G R L Q

3961
AAGTCTAATCTTACACGCGGAACAAAGGAATCTGAACCTTAAGAACTTAAGAAACCACTTGAAAATGCTGAAGGCTCAGAAATTCCTGCTGCCAGCGGGCTATTCAGGAGCACTGAATTTG
S L T M T A E Q K E S E L I E L R E T I E M L K A Q N S A A Q A A I Q G A L N G

5/124

Human genomic sequence

1 GATCAGACTT TGAAGAGTGT TTGTACCATG CTAAAGTTTA CAGAATTTAT
51 TCCTGCTCTT TGAGGGTGCA TTGCAAATCC AGGCTAGAGG GAGAGATACC
101 AGTTAGGAXA GTACAGCAAT ACTCTACTGG GAAATGGTGA GGTGTTTCGT
151 GAAGACAATG GCAACACAGA TGAAGACATG CAGATGGAGG AAATAAAGAT
201 CCAGTTGAGC TTGTTGGCCA GTTGGATAGA GGTGAGGTT ATGCATGATG
251 GAGCAATCTA GGTTTTTGTC TTGGGTAGGT GTTCCATGA TAGTACTCAG
301 AATGAATCAT ATAGTTGTAC AGGTTGAATC CCACCCATGT TTGCACAATA
351 GAGTGACTGT CTAGCTGAAA TCCAGATGAC ACTCTGTATG CTAAGCTATG
401 CTTCATGGAA CTGTATAAAG GCACTTGCTA CATAGGCTAG TGGCAGATCT
451 GGAAGTAACC TATATGGTAT ATAGGAAATG AGGTGGCTTT TGTATAAATC
501 CTACAGATAA ATTTTCATTTT CTGATCCTAT TATTTTGACT CATGTTAGCC
551 CAAGAAGAGT ATTCAGTACT TCATATCCCT GAAGGTAAGA CAGAGTAGTA
601 TTAGATTCAC TATTTGGCAA ATAAAAGGGA TCAAGTCCTA AGATCAAGCT
651 GATGAATCAA CACCTCATAG GATATGTCCC AACCAATTAT ATGGCTTCCC
701 CTATAAATAA AATCTAGTTC TCTTCTCTGG AGAGGAACAG TGAAGAATAT
751 CATAACCTAT GCTACAAACT GCTTGAGTAG GAGCTACTTC TCTCCAAGGC
801 TTTATATCAT TCATTCTGGC AGGCCCTCT GTTTGTCTC ACCAGCTCCT
851 GGGAAATTTA TTTCTCCTCT AGTGATATAA AAGCTCTCTG TTTGAGATGA
901 AGGGCTGCCC AGTTTATCAG ATCTGTATTA GTCTGTCTC AGGCTGCTAA
951 TAAAGACATA CCTGAGACTG AGTAATTTAT GAAGGAAAGA GGTTTAATTG
1001 ACTCACAGTT CCACATGGCT GGGGAGGCCT CACAATCATG GCGAAAGACT
1051 AATAAGGAGC AAAGTCACAT CTTACATGGC TGCAGACAAG AGAGCATGTG
1101 CAGGGGAACT GCTCTCCATA AAACCATCAG ATCTTGTGAG ACTTGTTTAC
1151 TATTACAAGA ACAACAGACA GGAAAACCCG CCCCTCAAT TCAATTACCT
1201 GCCACTGGGA CCCTCCCACA ACACATGGGG ATTATGAGAG CTACAATTCA
1251 AGATGAGATT TGGGTGGGGA TACCGCCAAA CCATATGAAG TTCTTTCTTT
1301 GTTACTGGGT ACCATATCCA TTCTGTTGAG GTTCTGAGCC TTCCAGTTA
1351 CTGTAACCTC TCTATCTCCT GTCTGTGCTA AGACTCAGTG ACCTCTCTCT
1401 GCCTTGCTTC TGCTTTGTCC TGACCCTTTC TGTGCATGCA CTCACTCTAG
1451 TTTGCCCACC TGAGGTGAGA GATGGTCCAG ATTAGCAACA ACAATCTGTG
1501 GACTAAAATC CTCTTTAGGG AGGAAGCAAA ATTCAGATGG ATGTTACTAA
1551 ACAAAGCTCA GAAACAGAGA CCAGGGTGTG GGAAGTAAGG TAGTAGCCTG
1601 AGAGCAGCTG GCAGTGTTTT AGACCTGGAG GGAGGTTAGG TCATCAGCAA
1651 TGAGGAGACT GCCTGGAAAA TCCTAGAAAA TTAAGACATC TGGTCAGGCA
1701 AGGTCATATC ACCAGCACAC TTCCCTTTTC AAGTTGAATC CCTTTCCTCT

Fig. 2

6/ 124

1751 GTTAAGAGGA TTCAAGTGTC TTTCTTGCAT TTTGTCTTCT CTTCTATATC
 1801 CATGCTTGCA ATATAAGGAG ACAGCAGTTG GCTGTTTGTG CTAGAAAATA
 1851 TAAATGGCCA TTTTGAAAGC ATGCCAGACA GGATCTGCGG CAAGTTTTCA
 1901 ATGTTACTGC TGCCATCTGT TGTTCCTCAG TGCTGGGATG TGAATCTCTT
 1951 GGCAAACATC TCTCTAATTC TGAACATATCT TTCACCCCCA TCTAGAGATA
 2001 TTCACTTACT GAAGTGCCTT TTTAAAGCAA TGTTCCTCAC CAAGGCGATG
 2051 TTCTGAATGT TTTAAAATGG AAGAATCTGG AATGTTTTTA TTATAATACA
 2101 TTTTGTATAT CCCAAAGCAA AAATCAATTT CTTCATGGTT AATACTTTTG
 2151 TAATTTTGTT TTTAATAATA TTTTCCTTTT AAATATAAGA AATATTTTAT
 2201 TGAATTAATA CTTTAATGTA GCTGTTTCAA GTAAGATAAA ACAGAACAGA
 2251 TTAAGTTTGT CAACCTTGTT CACAGTTAGC TCTGTAAC TAAGTTTGAG
 2301 CTTTATCTAA GCTTTTTTAT TTTTACATAA CGTTTCCCTT TTCACTTAAC
 2351 CTTGAAATTA TAGTAATTTG GGAACCTCTA TTCCTCTGAA AGAGAAAGCT
 2401 AATGCCAAAG ATATTCAAG GGAGAAAGAA GGTTTTTTAA AGGAGAGACA
 2451 ATTCAGCTCA GACTTAATAG CTGTGATTGC TATTTATTAA GCAGAACGCC
 2501 TATAACTAAA TTCTCAGATA TCCAAAAAAC AGCCTGTACA TTCTCAAAAG
 2551 TGAAGATTAC ACATTTTCTA AGTTAAGGTA AAAGTTTGT CTCTGTAGCA
 2601 TCTTACTGAT TTCTATCTTC TCATTCTGCC TTAATAATGT CACTAAATAA
 2651 ATGTTTGATG CACTAATACA TGAATAAAAC TATTCATGGT AATGATTCTT
 2701 TAGAAACACA GCTAAGTTTT GTAATTTTGT TTTTAAAAA TTAAAAATTT
 2751 AAATATAAAA ATGTTTTTAA AAGGCTTGAA TTTCTTGTTA AATGTACACA
 2801 TTTTAAGTTG TAGGCTGTCT TTAAAAATAA TCTCTCCACA CACTGTAGTA
 2851 TTTAAAACAT CATGATATTA CTATAAAACA TCAACAAATA GGGCAGTGGA
 2901 AAACATGGTA ATCACTAAAA ATGCTCACAT GTCATATATT AAGACTTGAT
 2951 AAGTAAACCA CAATAATAAA TAGAAAAGAA ATAGTTGTCT AAAAAGGGAT
 3001 TCTCACCTTT CAAACCTTAC CATAAAAATG GAATATAAAA GAAGGAAGAG
 3051 GAGGAGAAAT CAAATTATAT CATAAAATTT TCTGGGCAAA AATATTACAG
 3101 AAGAAAATAA GAAAGATTTA TGGAGTTGAC TGAAACATTT TTGAATCCTA
 3151 TACATAAAAA TATCGTTAAT TAAAAGGAAA AACAAAGAAA CAGATTTGGG
 3201 AAATATTTGA AACTGGTTTT TTTTTCATG TTAATAATGT AATACAAATG
 3251 GATTATTTAA ACTCCATTGC AAAAATACAC AAAGGACATT GACAATGTCT
 3301 GGAAATAAAA TTAGCTAAGT AAGTTATAGA AAAACTCAGT CTCACAATTT
 3351 GACAAATGTA ACTGAAAAT ATTAATATAA TTAGTAAC TAATTTACATG
 3401 TCAAAATTTT TGAATTACTA AAGGAAACCA CAATGCCTGA AAGTATCCAG
 3451 GGTTTTTTTT TTTTTTTATA ATATTGGCAC TGTCATATGG GTGGCAGGAA

Fig. 2 (cont'd 1)

7/124

3501 TTGAAGTGAT GTTGTCTTCT CAGTTATTAA GTTGCATCTG CAGTGTTTCA
 3551 AATGTCCAAA ACCTGTGAGT CAGTAATTCT CTTTTTGTAT ATTTATCCTA
 3601 ATACAATAAT TCTAAACATA ATCTCAATAT ATATGTACAA AGTTATTAC
 3651 TGCAGTGTTA CTTACAATAG TTAGAAAAAT GTAAAAATGCT TTATGCATCT
 3701 TAAAAATATAA ATTGTGTAAT ATATAATAGT CCATATGATA TAATTATATC
 3751 ATTATTATAA ATAATGAATT AGAAAAATAA TTAAGAGCAT TAAAAATAAT
 3801 ATAAGGTAAT ATGAAGTGAA TGAATAATGT ACAGATACTA TAATCAGCAG
 3851 AGTGTTAACT AGGTAAATTT TTATGTGTGT ATATACTACT TCCTAAAAAT
 3901 GACTTGACAG AAATCATCAA AATGCTAATG GTGGTTACTT CTGGGTGGGA
 3951 ATACAGATGA TTTACTTTGT TCCTTTTATG TATTTCTGCA CTGCCCAGTC
 4001 TTCCACAGTG AGCATATATT GGTTTTAAA TTTATATAAG ATGGAAAAAG
 4051 ATACCAAATG GTCTTCAATG AATCCTGGAG TTAACCTTCA TGTGTGTCAT
 4101 ATGTTATATT CTAACTTAT CACAAATAGA AGACTTTAAA TCAACTTGTA
 4151 CCTATTTCAA CTATATAACA GCATCTTTAA AATGAGCATT GAATTAACT
 4201 ACCAAAACCA ACCATCATGA GGATTATTCA AGTAATGTGT TTAAACAAAA
 4251 GAATTTGTAA TAAATTTACT TTATCTCCTT TGTGATTTCA GCCCATTTAA
 4301 AAAAAATAGA TGTTTCTACT CTCCTTCAGA TATCATTTAA ACATAAACTT
 4351 GTGCCTGACT GCATAAATCC CTTTTAACT AATATCACTT ATTACGTTTA
 4401 ACTAAGTCTA CCTAGGGCTT CCTTGATATA AGAACAAGAG CTTTCCATTT
 4451 TTTGTTTACC TAGCCCTTTC TGATGCCACG ACAGAATAGC TGTAATCTT
 4501 CATTATTTAT ATTCTAGAGA AAATAAAAGC AAATAAAAG GTCAGTGTAT
 4551 AAAGTTTATT GGTGTTCTC TTTACTCAA ACCCACATGG TATTAATGTT
 4601 AGTCTCTATG AATATTTTAT GGATAAAATC AGAGCATTAA GTGCATACTA
 4651 AAAACAATAA GAATGGAAAG ACTTTAACCT TATGTTTATA TGAATTTCTA
 4701 GGTATCAAG AAGTTTATAG GCTATAGGCT ATAAAGTCTT AGGCTATGAT
 4751 ATAGTAACCT AATGTAGACT TCCCTTGATA CATGAAAATA ATGGTACTAA
 4801 GTACAAACAG AAGATGAGCT TAAAATTATT CTTTGAGTCC TCTTGATGGA
 4851 TTTTTTCCCC CACACTTTCC CAAAATTTGT TTTATGCCTA TATTGTAGGA
 4901 GACCATGCAA GAGACCTAGA GTCTCTTTTT CTTTCATCAC TTTCCAATCA
 4951 ACAGCAAATC CTATCATTTT TACCACAAAA TATATCTTGA AACTCCCTTC
 5001 TTTTGATTTA CTTGTAACCT CCCATCAAAA ACTGAAGAGT GTCACAATAC
 5051 TTCATTAAAGT TCCCTACTTG CACTCTACCT TTAATATATT TGTAGCACTA
 5101 AAATGTTTTT AAAACATATA TCTGCTTATG TCATTTTACT GCTCAATACT
 5151 ATCTGATTTT CTATTGCACT TCTAAGATAC TCTAATTCT TAGCACTCTA
 5201 TATAAAATCC TTTAAGGGCT TCCCTGCTCA CCTTTTCAGA CTCAGAACTA
 5251 TGTATTTCTT TTTGCCTGCT GTACTTGATC CACTGGATTC TTGATTTTTG

Fig. 2 (cont'd 2)

8/124

5301 TTA~~CT~~TCCAG GTTTTTACAC TTATTTT~~T~~TAC AATAAATGTG AAATACCCTT
 5351 TTTGACAATA TCTACAAATA TTTCTTATTT GTCTTTATTG CTCTTTCCTG
 5401 TAATGTTTAG TCTTCATTTT CCTGATAATG GCTATCTAAA GTTATCTCCT
 5451 CAAAGAAGCA GTTATTTATT CACCCAAATC TTCTAGTCCT TCTCTGGAGT
 5501 TTTCTTCTCA CTTCATTCCC TTGGTTTTTG CCACAATTG TAATAATTTG
 5551 CAATTTGGAG TGTTAGAATG AGGGAATAAA TCACAGGTAA TGACTATAGT
 5601 TTGTGACTAT GTAAGATTGG ATTCGTTATT GATTTATTCC ACAAACACTG
 5651 AGGCACTGCA TTTAGCCAAA TGCCAATCTT GGGCAGTGAG ACTCTGAAAG
 5701 AGAATCTGCT TCCCCACCA TAAACTACAA AGTGAAACAA CTCAGAATGT
 5751 ACATAAATTA CAGAATGAAA GCACACTAGA AGTAAACACA GATGTGGAAG
 5801 AGGTAAAGTG TCCTTGAAAA TCATGGAAAG ATTCATAAAG GGAATGACAT
 5851 TTCAACTGGA TTCTAAACCA GTTATTCAAG CTCCACAAGG TTGCACAGTA
 5901 AATGAGCAGT GGCAGGATGA CATACCTTAG AAAGTAAAAG GAATCTTTTT
 5951 TAAACTGCTA TAAAAATCAT TACATATACA TTTTGTAGGT CGAGAGTAAG
 6001 GTATTTAACA TAAATCATT TTAGTATATC AGTGTTTATA TAGACTTAGG
 6051 TTTTCTCAT TTAAACCTC TTTTAATGAC TTGTGCTTTT CTTCATGGTA
 6101 ATAAACATT TTCCCAGGAA GTGCTGAATA AATCTTCTT GAAATACGTT
 6151 TTATTGCTTT CTATCAATGA CCCTGAAGTA ATACAGAATT TACACTTCAG
 6201 CGGTGCAAT GCTCAAAC~~TT~~ GACAGGTAAT GCACTGTGTT TGCTGATATA
 6251 AGAGGTATGA TGTAGGGCTA AGTGGTTTTG TGCTCATTTA GCTTTCAGGA
 6301 GAAAATAATT GACTTAACAT TTTGATACTA AAACCCAAAG CCTAACAGTT
 6351 AATCTTGGT ATTTTAAATT ATTATTGCAA AGATTATTGT GCCGAATAAT
 6401 ATGAAAATAT TTTATATAAT ATTTAAAAAG TATATCTCTT TCTTGGTATT
 6451 ATTTAAATTA CCATAAAAT GTGCGAAAAA GTTATACTGA AATGTGATAG
 6501 GATCTTTTAA AAGTGGTGCC TTGATTTTGT TAAGTGTTAC CTAGTTTTC
 6551 TCTGAAAACA AGAAACATAC CCAGAAGTTT TCACGAAATG GTCTCATGAA
 6601 TATCTAAGGT TAGTCCGTAG TCTCATCTGA GACAAGGAAA GTCCCTTCCA
 6651 CTATGAGCCT GTAAAATCAC AAGCAAGCTA GTTACTTCCT AGATACAATG
 6701 GGAGTACTGG TATTGGGTAA ACACAGCTGT TTCAAATGGG AGAAATTGGC
 6751 CAAAATTAAT GGGTTACAGG GCATGCAATT CCGAAATCCA TCTGGGCAGT
 6801 CAAATTGTAA AACTCCAAAA TGATXTCTTT TGA~~CT~~CCATG TXTCACATCC
 6851 AGGACATGCT GAXGCAAGAG ATAGGTTCCC ATAATCTTTG GCAGCTCTGC
 6901 CCCTGTGGCT TTGCAGGGTA TATCACCCT CCCAGCTGCT TTCACAGGCT
 6951 GGCATTGAGT GTCTGTGGCT TTCCCAGGAA CAAGGTGCAA GCTGTTGGT
 7001 GATCTACCAT TCTGGGGTTT GGAGGATGAT GGCCCTCTTC TCATAGCTCC

Fig. 2 (cont'd 3)

9/124

7051 ACTAGGCCGT GCTCCAGTAG AGACTCTGTG GGGGCTCTGA CCCCAGATTT
 7101 CCCTCCTGCA CTGCCCTAGC AGAGATTCTT CATGAGGGCC GTGCCCCTGC
 7151 AGAAAACTCT TTCCTGGGCA TCCAGGCATT TCCATACATC TGAAATCTAG
 7201 GTGGAGGTTC CCAAACCTCG ATTCTTAATT TCTGTGCACC TGCAGGCTCT
 7251 CTACCACGTG GAAGCTGCCA AGGTTTGGGG CTTGCACCCT CTGAAACCAC
 7301 AGGCTGAGCT ATACCTTGGC CCCTTTTAGC AATGGCTGGA GTGACTGGGA
 7351 CACAGGGCAC CAAGTCTCTA GGCTGCACAC AGTATGGGCA CCCTGGGCCC
 7401 AGCCCTCAAA ATCATTTTTTT CCTCCTAGGC TTC'TGGATCA GTGAAGGGTG
 7451 GGGCTGCCAT GAAGACCTAT GACATGCCCT GGAGACATTT TCCCCATTGT
 7501 CTTGGGGATT AACACTGGCT CTTGT'TACT TATGCAGATT TCTGCAGCCA
 7551 GCTGAATTTT TCCTCAAAAA ATGGGTTTTT CTTTTCTACT GCATTGTCAG
 7601 GCTGCAAAAT TTCTGAAC'TT TTATGCTGTT TCCCTTTTAA AATGCGATGC
 7651 TCTAACAACA CCCGTCACCT CTTGAATGCT TTGCTGCTTA GAAATTTCTT
 7701 CTGTCAGATA CCCTAAATCA TCTCTCTCAA GTTCAGAGTT CCACAAATCT
 7751 CTAGGGCAGG GGCAAAATGC CACCAGTCTC TTTGCTAAAA CATAACAAGA
 7801 GTCGCCTTTG CTCCAGT'TCT CAGCAAGTTC CTCATCTCCA TCCGAGACAA
 7851 CCTCAGCCTG GTCCTTATTG TTTATATCAC TATAAAAATT TTTGTCAAAG
 7901 CCATTCAACA AGTCTCTACT CCAAAC'TTTC CCACATTTTC CTGTCTTCTT
 7951 CTGAGCCCTC CAAATTGTTT CAGCCTCTGC CTGATACACA GTCCCAAAGT
 8001 TACTTCCACA TTTTTGGATA TCTTTTCAGC AATGCCCCGC TCTACTGGTA
 8051 CCAACTTACT TTGTTAGTCC GTTTTCACAC TGTGATAAA GACATACCCA
 8101 AGACTGGAAG GAAAAAAGG TTTAATTGGA CTTACAGTTC CACATGGCTA
 8151 GGGAGGCTTC ACAATCATGG CAGGAGGCAA AAGGCATTTT TTACATGATG
 8201 GCAGCAAGAG AAAATGAGGA AGATGCAAAC GCAGAAATCC CTGATAAAAC
 8251 CATCGGACCT TGTAAGACTT ATTCACTACC ACTAGGACAG TATGGGTGAT
 8301 ACCACCCCCA TGATTCAAAT GATCTCCAAC CAGGTGCCTC CCACAACACA
 8351 TGGGAATTAT GGGAATACAA TTCAAGATGA GATT'TGGGTA GGGACACAGA
 8401 GCCAAACTAT ATCACATGGA TTTCTTATAC TTTTGCTTTT AATAACACAA
 8451 ACAAAAAAAT ACATCATTAA AAGGTTAGAA GTGAGAAGGT GTTTTTATGG
 8501 AAATCAAAAA TAATATCACC TTAGTGAACA GTATTCTTAT GATTGTAGTT
 8551 GAATTAGAGA GCAGAATACA TCTAGAAGAT TCAGTAGTAA GCATGTTTCT
 8601 TCGATTAAATG GAAAATTTGA ATAGCCTAGC TGATTGAGAT TGAGGTTACT
 8651 ATTAAATGCC TGAAGTATAA GAGTTGGTTG TTTATGTAAA CAAAATATCT
 8701 GTTTTACATG TACATGTGTA AGTAGGACTG TTGAGCCCCA GTAACATGAA
 8751 ATATCAAGA GCATGACTCG AATACCTGCC ATATGAAGTG CTATTACATC
 8801 AAAAAAGAGG CGTGTGCTGA AAAATTACCT ACAAATGGCA TTTTCCTCAA

Fig. 2 (cont'd 4)

10/ 124

8851 ATCAATTTTA AATCTTCAGA ATTTTCATTTT AATAATTGTT TAGTTAATAT
8901 TTCAGAATCC CTCATCATAA AAAGCAGGCA AAAGGCAAAA GTCCTTGAAT
8951 GTATAACACA TTTGTTTTCA AACAAGCCTG CCTCTAACTG TGAATCCAGG
9001 AGTGAATCCA GAACTACAAA TTAACATAAGA TTGGCCCCAT CGAGTTACTG
9051 AACGTTAAAA ATCTAAAAAC TAAAAGGCAT GCCTCAACAA TTATTTTCTT
9101 CTTGGAATCA TTAATTAACC TATGTGTATC CAAACAATAA TCTTCCAGCA
9151 GTTTCGCTAG CTACATTTTT AATTACTTAA TATCATGTAA AATTTGTTTT
9201 ATTATTGTTC AGTTCTGAAT TTTGACATAT GCATCAAGCC ATGCAACTGC
9251 TACCACAGTC TTCCTGATCA CTGATCTGTT CTAAATCTCT ATAGCATTTT
9301 TCCTTTTCTT AAATGTTGCA TAAATAAAC CATACCTTAT GTGGCCTTTT
9351 GAATCTGGCA TCTTTAACTT AATGCGCTTG AAATTAATCT ATGTCATTTT
9401 ATGTATCAAT GGCTCAATCT TTTTAATTGT TAAGAAAAA TGTATGCTGG
9451 GATAAATATC TTTCTAAATG AGTTTGTGTT CACAATGCTG AGTGTGTTGTT
9501 TAGGATAGAG TCCTAGAAAT GGTATCACTA GGTCAAACAT TCAAATAATT
9551 TTAAAATATT TGATACATAT TGCCAAATAA TCTCAAATTT TTTACCAATA
9601 TACATTTATG ACAGTATGGG ATAAATGTGT CTTTCTTATA CCAACTGACA
9651 ACATTAATGA TAATACATAA AATATTCTTT GCTAATTTGA TGGGACAGAA
9701 ATGTTATATC CTTATTAGCA TTTTATTATT GTGGTTGAAT GACTGTACTG
9751 TACAGCCAGA GATATTTGGT TCAAAATCCA TCTTCATTAT TTACTGTATG
9801 TGAAAATTTA GGTGAGCTAT TTAATCTCTT GATGCCTTAG TCTCCTAATC
9851 TATAAAGTGG GGATAATTGT ACCAATCATA TTAGGTTCTT GTGAGAATTA
9901 ACTGAATTAC TATAGAAAAT GCTTAGAATG GTATCTAGTC ACCAGGAAGG
9951 ACTCTCTCTG TATTACTTGT TTATTATCTA ACACGTTTAA TTATTAATGA
10001 AGCTCAGTTT CGTTATATGC TTGGGATATT TGAAACTTTT CTTAGTGAAT
10051 TTTCCAATAA AATTATTTGT CTATTTTCTT ATGGACAAGT TGGTATTATT
10101 CTTACTGGTT TGTTTCAGGT TCAGTTAGTA AGAATTTTAA GGATTTTCTA
10151 TCACATTTTA GCAAACCTTT TCTGCATTTT ATCTTTTTTC TTTCAGATAA
10201 TGTTTGCAAA ATGTAAAAA AACAAAAGGT TTCTTCATCA AGTTGGTATC
10251 TTTATCTTTT TTATTGCTTT GTGATTTGAA AATTCTTGTC CTGAGAACCA
10301 AAATATATAT TTGATGAAAT AGTTCTCTTC TTTTACTCAT TCTGAAGTCA
10351 TTGGAATTGA ATTTGGCATA TGATATAAAT CCTAATTTTA TATTTTATGA
10401 TATTCAAAAT TTCTAACAAA TATTTACTTA ATAATCTAAT CCAGGTTTCT
10451 ATTGTTTCTT CTGTTTCCTT TATAATGCTT TTTCTGAAGT TATTTTTCCT
10501 AGACTTAAAT ATTAGTATAA TATTATCATA GAGGAAAAA TATCTGTTAG
10551 CTATGAATAA AAGGCTTTCA TCTTATTGTT GCATTAATAT ATTTAAATGT

Fig. 2 (cont'd 5)

11/ 124

10601 AGAGAGCATA CAGATTAGCA AAGAAAAAGT ATAATTGCCT TTTTATATAG
10651 TTGACATGAA CATGTATAAA GAAAAACCAA AAAAATCAAT AAAACAACATA
10701 GAACCTATTA GTGAATTTAG CAAGATCATA GCATACAAAG CCAAGATTCA
10751 AAATTCCATT TTATTTATCT ACTAACAAAA AATATTTGAA ATTTGAAAAAT
10801 TTAAATATGC CATTTACAAT AACATCAAAA TATTGAACAA TAAAGTATTT
10851 AGGAATTTAT AAAATGAAAT CTCCTATACC AGGAATTACA GACCATTGCT
10901 GAAATAAATG AAAGAAGACC AATATATGTG AAGAGATACT CATTTGTGGA
10951 TTGAGAGACA ATATTGTTAA AGTATCAGTA TTTCCCAAAT TAATCAATAG
11001 ATTCAATATA ATGGTGAACA GAACACCAGA AGATGTTCTG TCGAAGCTGA
11051 CAAGCTATTT CTATAATTCA AATGGAAATG CAAAAGGCAG TCACTGCCAA
11101 CACCAGCATG GACTGTCTGG GTTCCAGTAG GTTACTTCAC TACTGCCTCT
11151 TCTGTGAGCC ACATCAGCAG AGCTGCCCAG AAGCCAGAGA AACTCCTCAC
11201 ACCTGGCCCA CTGCTGCAGC TACCAGCATC CAGGCAAGCC ACCATCAGCC
11251 CACTGGTAAC TGCCAACAGA GGTACCACTG TACACTACCC TGGGGAACAA
11301 AGATAGGCAT GTAGTCAGCC CACCTCTGCC ACCACTAGGG CCTGAAGCCT
11351 GGCCACCTG ACACTGCAGT CCTCAGCACA GCTTCATCAC AGCTTCTGTT
11401 AATAACCACA CCCTAACCTA CCAAGGAAAT CACAAATGTC ACTGACACTG
11451 TTTGTAGCCA AAGAAATCAT AGAGAGACTA CATTACTGCA CACACCCATA
11501 ATCAAAGCCA CAGTACCCTA TCCAGACAAC ATCACAGGTA TATCTAAAGG
11551 AAAAAATTTT CCCATATGAA AGCGAATTCA AATATAGGAA GAAGCGACTG
11601 TTACAACAGA TATGCAGATA AAGCTTCAAC AATATCCTAC ATTCAACCAG
11651 AAGAAAGAAT CTCAGAAGGT AAAGACAGGT CTTCTGAAAT AATCTAGTCA
11701 GACAAAATTA AAAGAGAATA ATCAAATCCT TCCTGACATT TGGGATAACA
11751 TTAAAGTGAC CAAATATACG AATTATAGAT ACCCCTGAGA GTGAAAAGAC
11801 AAAGAAAAGA TTAGAAAACC CACTTAATTA AATAATATAT GAAAACTTCC
11851 TAAGTCTAGC AAGAGTTTGA GATATTTGGG ATGCAGGAGG CTCAATGGTC
11901 CCCAGGCCGA TAAAACGCAA AAAGGTCTTA TACACAGCAC ATTACAATCA
11951 GACTGTTTAA AGTCAAAGAT AAGGAATAAA TTCTAAAAAC AGCAAGAGAA
12001 AGTGTATGAT AACCTATGAA GTAAACCTTA TCAGACTGAC AGCAAATTTT
12051 TGGCAGAAAC TTTACAGGCC AGAAAGAATA GGACAATATA TTCAAAGTGC
12101 TTAAAGAAAA AAAAACTAT CAGCCTTAAA TACTATAGCC CACAAAATTA
12151 TCCTTCATAA ATGAAGGAGA AATAAAAGGT TTCCCAGACA CGAAAATGCT
12201 GAGGTAGTTT GTTACTACTA GACTGGACCT ACAATAAATG CTCAAGGGAG
12251 GTCTGGAAAC TGGTAGTGAA AGGACGACAT TTATCATCAT GAAAATACAT
12301 GAAAGTATAA AACTCCCTGG TAAGCAACTA AAGGGAGGTA TCAAATGTTA
12351 CCACCAGAGA AATCTAACTA ACCACAATGA CAAACAATAA GGGAAAAAGA

Fig. 2 (cont'd 6)

12/ 124

12401 AAGGAACAAA AATATATAAG ACAACAAATA AACAAACAATA TAACAGGAAG
12451 CCTCACATAT CAGTAATCAC TTTGAATGTA AATGAATTAC ATTCTCCACC
12501 TAAACGTTAT GAAATGCCTG AATGATAAAA CTATATGATC CAAATATATG
12551 CTGATTACAA GAAACTTACC AGGCAGACAT ACATAGGCTG AAAGTAAAG
12601 AATGGTAAAA GATATTCCTT GCAAATGGAA AGCAATAGTG AGCAGGAGTA
12651 GCTATACTTA AATTAGATCA TACAGACTTT AAGTCAAAAA GAGTAAATA
12701 AAAAAGACAA AGGATGTTAT TATATAATGA TGAGATTAAC CCAGCAATGG
12751 GAAATAACAA CTCTAAATGT ATATGCATTC AACACTAGAG AACTCAGATC
12801 CACAAAGCAA ATATTAGACC TAAAGAGAGA AATAGACTGC AATACAGTAA
12851 TAGTGGAGAA CTTCAACACT CCACTTTCAG TATTAGACAG ATAATCTAGG
12901 CAAAAAATCA ACCAGTAAAT TTTAGATTTA AACTAGATTT TAGACCAAAT
12951 GGACCTAACA GACATTTACA AAACATTCCA TCCAACCACT GCAAAATGAA
13001 ATTTGTGTCA TCAGCACATG AAACAATGTC CAAGATAGAC CACCATATGT
13051 TAGGCCACAA ATCATGTCTC AGCAATTTTT TAAAAGTTGA AATCATATCA
13101 CATATCTTCT CAGACCACTG TTGAATAATG CTAGAAATCA ATGCCAAGAA
13151 TAACGTGGGA AACTATACAA ATACATGCAG ATTAAACAAC ATGTTCTCTG
13201 TTGATCACTG GGACAATAAG GAAATTAAGC TGAAAATCAA AAAATTCTTG
13251 TAACAAATAA AGATTGAAAC ATAACATATC AAAACCAGTG GCATACAGCA
13301 AAAGCAGTGC TAAGAGGGAA GTTTATAGCA ATAAATGCTT AACTGAAAA
13351 AGTAGAAATA TTTTAAATTT AGCAACCTAA CAATGTGCCT GAAGAACTA
13401 AAAATCAAG AACAAATCAA ACCCAAAATC AGCAGAAGAA ACACAAAAAT
13451 AAAGATCAGA AAAGAACTAA ATCAAATAGA GACTAAAAAA ATACAAATGA
13501 TTAACAAAAC TAAAATTTGG TTATTCAACA AGATAAATAA AATTGATAAA
13551 CCGCTAGATA GACTAAACAA GGAAAAAGAA TATCCAAATA AACACAATCA
13601 AAAACGATAA AGGAGACATT ACAACAGATG CCACAGAAAT AAAAAGGATC
13651 ATCAGAGACT ATTATTAACA ACTATATGCT GAAAAATGGA AAATATAGAG
13701 AAATAGATAA ATTCCTAGAA ACTTACAACC TACCAAGCTG TTGCATCAGG
13751 AAGAAATAGA AAACCTGAAC ATATCAGTAA TGATTAGCAA AATTGAATCA
13801 GTAATAAAAA ACATCTCCCA ACTCTTTTAA AGCTTTGGAC CAAATAGCAT
13851 CACAGCCTAA TTCTACCAAT CATGCAAAGA AGAATACCAG TCTTCTTGAT
13901 GCTATTACAA TAAATCAGAG GAAGGAATTC TCTCTGGCTC ATTCTACATG
13951 ACCAGTGTCA CCTTGAAACC AAAACCTGAC AAGGACACCA CAAAAAGAAA
14001 ACTACAGGCC AATAACCATG ATGAACACAG ATGCAAAAAT CATTAACAAA
14051 ATACTGGCAA ACGGAATCCA ACAGCACATC AAAAAAATAA TATACCACAA
14101 TCCAGAGGGT TTGTATCAAG GATACAAGTA TGAATCAATG TAAATAAATC

Fig. 2 (cont'd 7)

14151 AATAAACATG ATAAGCATCT TCACAGAATA TAAGACAAAT GAATATATGA
14201 TCATCTCAAT AGATGCAGAA AAAAATTTTT GATAAATTTT AACATCTCTT
14251 CATGAAAAAA ATCTCTAAAA CTCAGCATAG AAGAAACATA CCTCAATATA
14301 ATAAAGGCCA TATGTGACAA ACTCAGAGCT AATATCATAC AGAATGGGGC
14351 AAAGTTTAAA GACTTTCCTC TAAGAACTGG AACAAGACAA GGATGCAAAC
14401 TCTCACCCTT CCTATCCACA TAGTACTAGA AGTCCTAGCC AAAACAATCA
14451 GACAAGCAAA AGAAATAAAA AGTATCTAAA TTGAGAAGAG CAAGTAACAT
14501 TGTTCCTCTT TGCTGATGAT ATGGTTTTGT ATCTGGAAAA TACTAAAAAC
14551 TCCAGCAAAA ACCTCTTAGA TTTGATTAAT TAATTTAGTA AAGTTTCAGG
14601 ATACAAAATA AAAATACAAA AGTCAGTAGC ATTTCTATGC CCAATAATA
14651 AAATAGCTAG GAAAGAAATC AAGAAAGTGA TCCCATTAA ATTAGCTACA
14701 AAAAATTAAA ATACCTGGGA ATAAATCAAG GAAGTTAAAG ATCTCTGCAC
14751 AAAACTACAA AACACTGATG AAAGAAATTA AGGATTAAAC AAACAAATTG
14801 AGAAACATCC CATGTTTATG GATCAAAAGA ATTAATATCA TTAAATGAC
14851 CATACTTCCC AAAGCAATTT CCACATTCAA TGCAATTTCT ACCAAATTAC
14901 CAATGTCATA TTTCATAGAA TTAGAATAAT CCTAAAATTA GTATGGAATG
14951 AGAACAGAGC CCAAATAGCC AAAGCAATTC TGAACATAAA GAACAAATCT
15001 GGTCTGACT TAATCACTAT GCAATCTATG CATGTAACAA AATTGAACAT
15051 GGATTTTATC AATTTGTACA AATAAAAAAA TGTAAAAAAA GAACAAAGCT
15101 GGAGGCTATA GTAGCCAAAA CAGCATGGTA TTTTGTAGACA AATGGAATGG
15151 AATAGAAAGC TCAGAAATAA AGCCATATAT ATATATTGTG TGTGTGTGTG
15201 TGTGTATACA CACATACATG TATATATAAT GTGTACATAT AATGTTTCT
15251 ACATGTTCTA ATATTTATAT TCCATTCCAT TATACATATT CCATTTCTGT
15301 ATATAGGTTA TATAGAATTG GAAGACTATC TGCCATTAAA AAGATGAAA
15351 TCCTGTGATT TGCAGCAACA TGGTTGAAAC TGGAGTTCAT TATCTTAAGT
15401 GAAATAATCT AGGCACAAAA AGATAAATAT CACATGTTCT CACTTATATG
15451 TGGGAGCTAA TAACCTGATT ACATGAAGGT GGAGAATGGA AAGGTAGGTA
15501 GGAAACAGAG ACTGGAAAGG ATGAATGGAG GGTAGGAGG AAGGTGAAGA
15551 GAAGAGAGTT AAAAGGTGTA AACATATAGT TAAAGAAAT AAATCAATG
15601 CTTGATAGCA GAGTACAGTG ACTACAGTTA ACAAATGTA TTATACTCAG
15651 GTGATGAACA CCTAAATACT TGATCACTAT GCAATTATAT ACGTGAACA
15701 AAATCACTAT GCACTATATA CGTGTAATAT TAAATGCGTA CAAATAAAAA
15751 TAATAAAATA CTAATCCAGT ATCATTCCT GACAATGTTA ACTCAGGTGG
15801 ATAGGCATTA AGTCAATACT ACTATAAGAA CCACTTCTTG TTTATGTTAA
15851 TGCCATATAG AATGAAATAA AATTCATAA AATCCAAAA ATTAGAAAAA
15901 CTATCAAAAC TCAATAATAT TAAGACAACC CAATAAAAA GTGGTCAAAG

14/ 124

15951 GATTTGAACA TACATGTCAC CAAAAAATAT ATTCAAATTT CCAATAAATA
 16001 CATGTAACAA TGTTGACAT CGTTAGTCAT CAGAGAAATA CAAAAATAAA
 16051 TGGTAATGAG ATACTACTAG ATAGGCTTTT ACAGAGACTG ACAATACCAA
 16101 GTATTGACAA GGATATGGAG CAACTGAAAT TCTCATTCCT TGTGGTAAGA
 16151 ATGTACAATT ATATAACCAC ATTGAAAAAA CAAGTTTTCA GTTCTTTTAT
 16201 TCACCCAAAA TATATGTCTT TTGGAAAAAA TTTTTCAG TCTGTGGGTT
 16251 GTCTTCTCAT TCTCTTGATA TATGTCTTTT CAAAGAGGCT GAGCTTTACT
 16301 TTAGACAGTG GTCATCAAAG TGTGTATATT TGTGTTTTTA TAATTTATAT
 16351 GCATATATTC CTGTGAAAAG ATACTGTATG CATTGTTCAA CATGTACAAA
 16401 TATAAGAAAG ATATAGTAAA GAAATATATA TTTCTAAAT TATAAATGTA
 16451 TTTATTGGTG TTCCACGTTG CAAACTAAAT AATCTACGTT GGCTAATTTA
 16501 AGGAATTAAA CTATAGTAGA AGGTTCTCAT TTATTGGGAT GATTAGAACC
 16551 AGCCTTTTTG CAGGCTATTA GCGAATCATA GCACTAGGGC TTCACTGCTA
 16601 CCTCCACTGA CACCTCTGAC ACTTGAAACT TGAGGCCAGA TATCTGCCCA
 16651 TGCTGATAGA AAACAACCTGA ATAATTTAAT TTGCTAGATA ATAGAAAAGA
 16701 ATCAAATGAC TCTGCCACAT TGCTTGCCAG AAGATTGTTT TTCTCATTTG
 16751 TGACCTCTTG CCTATAAATG ATAGATAGTC CCTGTGCTGC ATGCTATAGG
 16801 TGTTGTAAG AGAGTCTGGG AATGTGAGCT TTTTATATCC TATTTTGGG
 16851 TGGTAAAGGT CATTCTATTA GTCTGTTCTT AAAGTGCTAA TGAAGACATA
 16901 CCCCAAATTG GGTACTTTAT GAAAGAAAGA GGTTTAATTG ACTCACAGTT
 16951 CAACATGACT GGGGAGGCCT AAGGAAAGTT ATAATCATGG GGGAAAGGGA
 17001 AGCACACATG TCCTTCACAT GGTCAGCAGG AGGATAATGA GTAAAAGGGG
 17051 GAAAAGCCCC TTATAAACT ATCAAATCCC ATGAGAACTC ACTCTCACAA
 17101 GAACACAATT AGAGTAACTG CCCCATGAC TCAATTACTT CCCACCAGGT
 17151 CCCTCCCACA ACACATGGGG CTTATGGGAA CTACAATTCA AGATGAGATT
 17201 TGGGTGGGGA CACAGCCACA CCATTTTATT CCACCTCTGA CCCCTCCCAA
 17251 ATCTCGTGTT CTCACAATTC AAATACAATC ATGCCCTTCC AACAGTCCCC
 17301 CCAAAGTCTT AACACATTTT AGTATTAACA CAAAAGTCCA AGTCCAAAGT
 17351 CTAATCTGAG ACAAGGCAAG TCCCTTCTGC CTATGAGCCT GTAAATTCGA
 17401 AAGCAAGTTA GCTACTTCCT AGATACAATA GGGTCACAGT CATTGGGTAA
 17451 ATACACACAT TCCAAACGGG AGGAATTGAC CAAAACCAAG GGGCTACAGG
 17501 CCTCATGGAG GTCCAAAATC CAATAGGGCC ATTGTTAAAC CTTAAAGTTT
 17551 CAAAATTATC TCCTTTGACT TCATATCTCA CGTCTAGGTC ATGATTATGC
 17601 AAGAGGTGGG CTCCACAGC TTTGGGCAGC TCTGCCTCTG TGGCTTTGCA
 17651 GGGTACAGCC CCACTCCAGG CTGCTTTTAC AAGCTAGTGT TGAGTGCCTG

Fig. 2 (cont'd 9)

15/124

17701 CAGCTTTTCC AGGCACATGG GTGCAAGCTG TAGGTGGATC TACCATTCTG
17751 TGGTCTGGAG GATGGTGGCC TTCATCTCAC AGATCCACTA GGCAGTACCC
17801 CAGTGGGGAC TCTGTGTGGG GGCTCTGATC CCACATTTCC CTTCCACACT
17851 GCCCTAGCAG AGGTTACCA TGAGGGCTCC ACCCCTGCAG CAAACTTCTG
17901 CCTGAACATC CAAGCATTTT CTTACATCCT CTGGAATCTA GCGGAGGTT
17951 TCCAGACCTC AATTGTTGAC TTCTCTGCAA ATGTAGGCTC AACACCCCAT
18001 GGAAGCTGGC AAAGCTTGGG GCTTTCACCT TCTGAAGCCA TGGCCTTAGC
18051 TGTACCTTGG CCCTTATTAG TTAAAGCTGG AGCAGCTGGG TTGCAGGGCA
18101 CCAAGTCCCT ATGGTGCATA CAGCAGGGGG GCCCTGGACC CAGCCCACAA
18151 AACCAATTTT CCCTCCTAGG CTTCTGGGCC TCGATGAGT AGGGTTGCCA
18201 CAAACTGTC TGACATGCCT TGGAGACATT TTCCCTATTG TCTTATTAAG
18251 ATTTGGCTCA TAGTTACTTA TGCAAATTTT TGCAGCAGGC TTGAATTTCT
18301 CCTCAGAAAA TGAGTTTTTC TTTTCTATGG CATCATCAGG TTGCAAATTT
18351 TTAAACTTTT TATGCTCTGC TTCCCTTTTA CAATTAAGTT CCAATTCCAA
18401 ACCATATCTT TCTGGATACA TAAACTGAA TGCTTATAAC AGCACCCAAA
18451 TCATATCCTG AACACTTTGC TTCTCAGAAA TATCTTCTAC CAGATACCCT
18501 AAATTATCGC TCTCAAGTTC AAAGTACCAC AGATCTCTAG GGCAGGGGCA
18551 AAATGCCACC AGTCTCTTTG CTAAAGCATA ACAAGAGTCA CCTTTGCTCC
18601 AGTTCCCAAC AAGTTCCTCA TCTCCATCTG AGACCACCTT AGCCTGGATT
18651 TCATTGTCCA TATCATTATC AGCATGTTGG TCAAAGCCAT TCAACAAGTC
18701 TCTAGGAAGT TTCAAACCTT CCCACATCTT CCTATCTTTT TCTGAGGCCT
18751 CCAAAGTGT CCAACTTCTG CCTGTTACCC AGTTGCAAAG TTACTGCCAC
18801 ATTTCTGGGT ATCTTTACAG CAGTGCCCCA CTCCTGGTAC CAATTTACCA
18851 TATCCATTTA TTCTCATGCT GATAATAAAG ACATACCCAA GGCTGGGTAG
18901 TTTATAAAGA AAAAAGAGGT TTAATTGACT CACAGTTCAG CATGGTTGGC
18951 AAGGCCTCAG GAAACAGAAT CATGGTGGAA GGAAGCAAA CACATCCTCC
19001 TTCACATGGT GGCAGGGAGA AGAATGAGCA AAACGGGGGA AAAACCCTTA
19051 TAAATCATC AGATCTCATG AGAACTCACT CTCTTGAGAA CAGCATGAGG
19101 GTAACCATGT CCATGATTCC ATTACCTCCC AACGGGTTC TCCCATGACA
19151 CGTGAAGATT ATGGGAACTA CTACAATTCA AGAGGAGATT TGGGTGGGGA
19201 CACAGCCAAA CCATGTCAGT CATGATATGA GAAATTATCA AATTAAGATG
19251 TAGGGAAGGT TTTTAAAAGA TTTGAGCAAC CACAAATGAC AGATATGTGC
19301 TATAGTAGTG CAAAATACCA TTTTGCTCTT ATTAATAATA TAATTGTTCT
19351 TGATAATCTG AATTATAAAT GTCATGGATA ATTATGATGC ATTATGCTCT
19401 CAGCAGCTAA AACTTCAAGC AAAATACACA CCTAGAGAGC AATCAGCCTT
19451 AACAATAATT CTATAAATTT AATTTTCTTT ATTTCTGATA ATTACATTTT

Fig. 2 (cont'd 10)

16/124

19501 AGTTGACTTC ATATGTGATC TAAATACATT ACCATTATTT TGGACTTATG
19551 ATGTAGCTCT TGAAGTACAT ATATGATGTA GCTCTTAAAG TACATATAGA
19601 AGAGCAGATA AAGTATCAGT TCACCATTTT TTTGTAGTTT GTGCTTTCAT
19651 GATGAATATT CTCATCAATG TACAGATTAT TTGCAGGAGC CTTTTAAATC
19701 CATGTGTCCA TTTTATGAGA CTTAGCTTTT GTCTGTATAT AATGTGTTTA
19751 TTCAGTGTGC ATGGATTAAT TTGAGAGAGC ACAGGTATGG GTATCTTTAC
19801 AGCAGTGCCC CACTCCTGGC ACCAATTTAC TGTATTAGTT TATTCTCATG
19851 CTACTAATAA AGACTATATA TCACAATAAA CTGAGAACCA GCTGGTAAAT
19901 GAGAGAACTG TGGTCCACCT TTTCATTGTG GAGTTCTCAT TTTCTTAGC
19951 TTATGCTGCT TATTCAACAC TATTTCTGCA TAATCTAATG CATTCACTAA
20001 ATGAAGGTGC TGTGTTAGCC TCCACATGAT ATTAATACAG CCTATTTAAT
20051 TTATCCTTCT TTAGATTAAA AATAAATAAG TAGTCATGTG CCACAGAATG
20101 ACACCTCAGT CATTTGGTCA TTGAAGGACC ACATCTATTA CTGTGGTCCA
20151 ATAAGATTAT AATAACATAT TTTTCTGTG CATTTCATT GTTCTGATAT
20201 GTTTTGATAC ATAAATGCTT ACCATCGTGT TAGAGTTGCC TGCAGTATTC
20251 AGTACAGTAA CATGCTGTAC ACCTAGGAGC AACAGGCTAT ACCACATACC
20301 TTAGGTGTAT AGTTAGGTTA TACCATCTAG GTTTGTATAA GTACACTCTA
20351 TGATGTTCTC ACAATGAACA AAATCACCTA ATGATGCATT TCTCAAAACA
20401 TGTCCCTGTC ATTAATACAG TATGTAACAA TACAGTTAGT ACAATATGTA
20451 ATACATGACT ATATTCAGAA TTTTAGCTAT TTCTCTTATA TTTCAAATGG
20501 ATTTTCTTAT GCACTGTGTG GCACGGGCAT TTCATTTTAG TAACCACAGT
20551 CTGGGAAAGG AGAAGTCTTT GAAGGATGTT GAGCAAGGTT ATGACATGGC
20601 CAGATGTGAA TTTTGTATCA GTGACTCCAT GTTAGCAGAT AAAGTTGTAT
20651 TGGGAAAGAT CAAAAGCATG AAGGCCAGAT AAGAGGATAC TGTATGTTAT
20701 CATGGATGGA AATGTGAGGG ATGGCAGGAG AGATGCTATG ATTGAATGAA
20751 TCTCAATATT CTTGGTGATC AAAGAATAAT GAGACTCATC CAATAAGACT
20801 CTGTGAATGA TTGAATGTAG TTCCTAAGCT AGGAGGAAGA ATGAGGAATG
20851 ATTTTCTGGT TCCTGACTAC AGCACAAGTT TTTGATTTTT AGAACAAAGA
20901 ATAAATTTGT ACATGCTTTA TGATTCCTGG TTGAATTTTT AAGGATAAAA
20951 AAGTCAGCTG TAATATTATT CTTTCCTGAT ACCATGCAGT ATTTGTATCA
21001 GTGATCTTAT TCATTCCACA CACATTCTTC TTGAACCTGG AACTGCTCT
21051 AGACACTGAT TCTTTCCAAA TATCAGATAA GGTTATTCTT ACGTAGACCC
21101 TCAGTTCATA TAAATATGAT TTTCCCAAAA TGTGAAATAA GTGACTTTTC
21151 ATAAGATATT TTTTAAAAGA ATGTCTTAAT AATAAATTGT GAATGTTGCA
21201 TGGAAATGTA GGTGACTTGC ATTGTGCATC CTGTGTTTGA TTTACTGCTC

Fig. 2 (cont'd 11)

17/124

21251 TTGCATGTCT TGCCTTTAGC TGGGATGACA GCAGTTCAGT GAGCAGTGGT
21301 CTCAGTGACA CCCTTGATAA CATCAGCACT GATGACCTGA ACACCACATC
21351 CTCTGTCAGC TCTTACTCCA ACATCACCGT CCCCTCTAGG AAGAATACTC
21401 AGGTGAGAAT TACCACCTTT CTTTTTCCAG TGTTCCTGCC AGCTTTTTTCC
21451 CCAAAATTAC TTAATATTAG ATTAAGGTAT AGCACAAGCC CTTAATCCAA
21501 AATTATTACA GAAACTGGAA AATGCAGAGA TAATAAGGAC TCCCTTTGCC
21551 ACTCCTGAAC CCTGAAGCAT CTTTCATCTT AGTCTTTCCT AAAGCCACAA
21601 CCCTTAGGAG GAGCAACAAT GTGCACTGCA GCCAATTTTG AATAAACAGA
21651 AGCAGCTTAT ATATATATAT ATATATATAT ATATATATAT ATATATGATA
21701 TACATTACAT ATTTATATAT ATGTAATATA TGTGCCATAT AGCCTGGTGG
21751 TATAGTTATC TATACAAATA TATTTATTTA TTGTTAATAT ATAGAGTATA
21801 TAAATATCTA TTTATATAAT AGATATTTAT ATATATTAAA TATCTATTTA
21851 TATAATAGAT ATTTATATAT ATTAAATATA TAAAAATATA TAACATATAA
21901 TAGATATATA TTTTATATAT TATATAAATA TATATTTATA TATTTAATAT
21951 ATTAATGATG AATTACTATA TTTGTATAGA TAACTACACC ACCAAGCTAT
22001 ATGGTGTGTA TATATTAATA TATAATGTAT AATTCTATAT TAATATAATA
22051 GTAACATATC AATACTTAAT ATAATATATA TTCAATTGAT TACAATCTAA
22101 TTCAGAAAGA TTTATGTTGC CATATCTCTC CTTACAATAT CGATATGTTT
22151 GTTTAAAAAT CCAGCAATTA TTTTCATAGT CTAATTTTAG ATAGTTCTTG
22201 ATTAATTTTA TATGATCTCT GAAATATATC ACTGGATCTG TTGTGAATGA
22251 TAAATCAAAA ATGAAAAATG GACATTACAT CATTAAGTTC TAGCTTGTCT
22301 TACTACTTCT TATGACATTT GATATAGAAA ATTTCTACCT TTCTGTAGCG
22351 TTTAATTGGT GTTTTCTGCA TGTATTTATT CTGAAATTCT CTAATATCTG
22401 CAAGTGGGAA TTATGTGGCT AAAATTAATA AAATGTAAGT GAAGGTAAAT
22451 CAAAATAGAA TCTTTGGATT TATCCAGTTA TCTGAAAGTA CATTTTCATTG
22501 CCTTAATTCA CACTTTTATAA ATTTTCTAC ATAAAGTTTT TCTGTAATAT
22551 TTGTCTTTAT AGCTGAGGAC AGATTCAGAG AAACGCTCCA CCACAGACGA
22601 GACCTGGGAT AGTCCTGAGG AACTGAAAAA ACCAGAAGAA GATTTTGACA
22651 GCCATGGGGA TGCTGGTGGC AAGTGAAGA CTGTGTCCTC TGGACTTCCT
22701 GAAGACCCCG AGAAGGCAGG GCAGAAAGCT TCCCTGTCTG TTTCACAGAC
22751 AGGTTCCTGG AGAAGAGGCA TGTCTGCCCCA AGGAGGGGCG CCATCTAGGC
22801 AGAAAGCTGG AACAAGTGCA CTCAAACAC CCGGTAGGCT TGTGTTTTGC
22851 CAGCTGTTAT GCAAAAGTGC TTTACTTTAT TGTTTCCATT CAATCTTTGT
22901 TTTCTCTAAC AATAGCATTT CTAAAATACC AAATCTTTAT CCATATTAAA
22951 CATGGAGTCA AATAGTTAAA TAGTTTTTCT GTCTACGTTT CACAACTCG
23001 TCATAGAAGC CCAAGTAGGG CCTATATCTA GGCATTCTCT GGAAAGCCTC

Fig. 2 (cont'd 12)

18/124

23051 CTCATAAACT AGGGGTACTG GATGCCTTAC CTTGCCAGAG TTATTTTCAGG
 23101 TAATGGGGAA ATAAGATTAG GTTGCTAAAG CAACAGTTAA GTTTTTTTGT
 23151 TTTTGTCTG CGTTCCTAAT GAAAGTTTGG AATTTTACAC CTAAATATGC
 23201 CACTGAATTG CACTACAGAC TCTGAGAGGA ACAAGCAATG ACACTAATCA
 23251 ATTGGAATGC TGGAGATTTG AAATATTGTC TGTGTATTAG ACTTCATGAA
 23301 AGAAGAGAAT GAAATAGTTC TTCAAAATTG TGCCATACTT TTTTAAAAA
 23351 GACTCTCCCC GTATTTTAA AATAATGCCT AATTATAAAT AGTGCCACCT
 23401 GAAGCACTAA TTAACAGGGT ACTCCAAATA TAATCATCTC ACAGATATTC
 23451 AAATGAATTC TTTTCTAGT AATTAGCTTG ATAGGGTTAA GTGTTACCTT
 23501 TTTAAAAAGA GTTGCAAAAT ATAAGACATT AACAAATAGC AAAACATATG
 23551 TTTTCATTTT ATCTCTTCCA TCTCTCATAA TGTTTCTTCT GACAGCCAAA
 23601 TTTTGTAGC TATGCACTCA GTCCTCTCAA TATATGAGAT TTTTGATCTA
 23651 AGCCAATACA TTTAGGAAGG GAAATAATAT AAAGAAGCAT TCACATTTTA
 23701 CACATTGTTT CACGAAGTGT GGTGATATCA AACTCTACAG GCACATATAT
 23751 TTGTGTATTT CTCCTTAATT AGGGAAAACC GATGATGCCA AAGCTTCTGA
 23801 GAAAGGAAAA GCTCCCCTAA AAGGATCATC TCTACAAAGA TCTCCTTCAG
 23851 ATGCAGGAAA AAGCAGTGGA GATGAAGGGA AAAAGCCCCC CTCAGGCATT
 23901 GGAAGATCGA CTGCCACCAG CTCCTTTGGC TTTAAGAAAC CAAGTGGAGT
 23951 AGGGTCATCT GCCATGATCA CCAGCAGTGG AGCAACCATA ACAAGTGGCT
 24001 CTGCAACACT GGGTAAAATT CCAAAATCTG CTGCCATTGG CGGGAAGTCA
 24051 AATGCAGGGA GAAAAACCAG TTTGGACGGT TCACAGAATC AGGATGATGT
 24101 TGTGCTGCAT GTTAGCTCAA AGACTACCCT ACAATATCGC AGCTTGCCCC
 24151 GCCCTTCAAA ATCCAGCACC AGTGGCATTG CTGGCCGAGG AGGCCACAGA
 24201 TCCAGTACCA GCAGTATTGA TTCCAACGTC AGCAGCAAGT CTGCTGGGGC
 24251 CACCACCTCG AAAGTGAGAG AACCAACTAA AATTGGGTCA GGGCGCTCGA
 24301 GTCCTGTCAC CGTCAACCAA ACAGACAAGG AAAAGGAAAA AGTAGCAGTC
 24351 TCAGATTCAG AAAGTGTTTC TTTGTCAGGT TCCCCCAAAT CCAGCCCCAC
 24401 CTCTGCCAGC GCCTGTGGTG CACAAGGTCT CAGGCAGCCA GGATCCAAGT
 24451 ATCCAGATAT TGCCCTACCC ACATTTTCGAA GGTAAGGATG TATAAAATGA
 24501 TGCTGGAAAA ATATAAAGGA TAAATATGTG TTAGACACAT ACATTACATA
 24551 TAAATGTGTG TATATATATA TTTTAAATAT GTATAAGGTA TATAATATAT
 24601 ATATCTTAGA ATTCTTTAAA GTACACAGTG AGCTCTATGA AGCTTATCAT
 24651 ATAAACAGCT AGCAAAAAAA ATAGTTCTCA TTTTGAGAAA CAGTCAAAT
 24701 TCAAAGTTTC ACTGTCATTG TGATACTAGC AACACAAACA TCTAAGAGAC
 24751 TTAAAAGCTG ATGGTAATAC CTAAGTGTAG TGATAAGGCA AAGTAATAGC

Fig. 2 (cont'd 13)

19/124

24801 TTGTAAAATT TCTATAGATT TCCATTCCCTC CTTTTCACAT TAAAAATTAA
24851 AACCAAATAG GTTTTCATGA CTTTGGCAT TCATTCCAG TGTCATTTTC
24901 TTGCTGGCTC TTAATGAGTT GGTGATCATA AATGTAGATG AAGTTGTTTT
24951 CCTTGTAACA GATTCCATTG GACAGATTTA TACAGTGTC AATCTTGACA
25001 CATTAAAGAC AATCAAGATA TGACATAATT TGAACTATT CCAGTGTTTG
25051 GTACAGTATC ACAACTGAAG AGTGGGCTAA GCTTCTAAC TCTTCATCTG
25101 CTTTCTTTGA CATGACTCTG GTAAGGATCA TGACTTGGTT TCTGTTCCCTG
25151 GATTGTTTTT GGTGTTAAAT ATGTGAAGTT CTGCTCTAAG ATATCACTGT
25201 TTTTAAATAC CCATGTGTTT TTAAGTGGTA GGAAAATAAA TGCAGTTAAA
25251 AATTGGGGAC AAATATCTAA ACCTCTCTGA GTCTGTTTTT TCATCTGCAA
25301 AATGGTAGAG TGTGGTTTAT AGTTCATTAT GGGTTCAATA TTTTAAATGT
25351 TTGTTTTTAT TCTGTTGACT AAACCCAGAA CTTTGATATC TTGGAAAGGA
25401 AAGATTTTGA AACATTTATT TTACAATAAA GCAATTTTCAG ATACCTGATT
25451 GTTTGAAAAA CCTAAAGGCT TTATTCCCTCC GTAGTAATAT TAATGCTGCA
25501 GAACTGTCTT TTTAAAATAC TGATTCTCAT TGGGAAGAAT GAATTATGGC
25551 GTATAGGGAG AGTAAATATT TCTGTTTCTT AAGTAAAGC CAATAGTGCC
25601 CTCCTGTGGC CCATTACCTA TGAAACAATT TCTCATATTC GTCATAAAAT
25651 ATTTCACTGT AGGAAATATG GATTTTCATTG CAACTCAATT AGTAATCATT
25701 ATGCCATTAC TTCATATCAT TGTATTTCCA TATTTACATA AATTTGATTC
25751 TACCATCTGC TTCATTTACA AAATAAAAT GTTTTCTGAA CTAAACTCCA
25801 AAATCTAACA GCACCAGCTC TGTTTCAAAT CACTATTAAA AAATGTATTT
25851 GAATAGCACT GGCAACTGAC ATAAAACCCCT TTGGCCTCTG CTGGGGAAAA
25901 TACAGACAAA CTGACTTGTT GCCGACAATA TCAATATTGT TTCCAACCAA
25951 CTGCTCCCTG ACAGTGA CTC AGACCACCAG ATACTCAACA CAACTCCCTA
26001 AACTTGCTTT AAGCGTTCCA TCTAGATTTT GAATAAACTG TTTAAAAATT
26051 TAAAAATAAA AAAAAAGAG AAGAGCTCAT TTAAGTGTG TCTATCGAAT
26101 GCGTAGAAGT TGTTTCATTA TAATGGTTCT GTAAATAGGT AACAGCAAGT
26151 ATGGTCAAAC TACTGACTTT GAGTGAAAGT CTCATGATCA CTTAAATTAT
26201 GAAAACCAGG GGTTTTCATG TTTGACTTAC TTTTGTTCCA CCCACTTCCC
26251 CTCTTTCCCT AGTAGCAGCT CAGTACTGAC CTACCCTTAT ATGAGAGATT
26301 TTCTGCACTT GATAAAGAAG TCCAAGCTTA TAAAAGTTCA TTAACATAGA
26351 GACAGGAAGT GCTTTGTAGT TCAGTACATC AAAGCACACT TGGCTCTGTG
26401 TACTGTAACC CGAAATATTA AATGTGGATA TTAGCTTCTT GGAACAACCTG
26451 AAGTTGTTAT TTGTTTTTCT TTTAGGTTGT TTGGTGCCAA GGCAGGTGGC
26501 AAATCTGCCT CTGCACCTAA TACTGAGGGT GTGAAATCTT CCTCAGTAAT
26551 GCCCAGCCCT AGTACCACAT TAGCGCGCA AGGCAGTCTG GAGTCACCGT

Fig. 2 (cont'd 14)

20/124

26601 CGTCCGGTAC GGGCAGCATG GGCAGTGCTG GTGGGCTAAG CGGCAGCAGC
26651 AGCCCTCTCT TCAATAAACC CTCAGACTTA ACTACAGATG TTATAAGCTT
26701 AAGTCACTCG TTGGCCTCCA GCCCAGCATC GGTTCACCTT TTCACATCAG
26751 GTGGTCTCGT GTGGGCTGCC AATATGAGCA GTTCTCTGTC AGGCAGCAAG
26801 GATACTCCGA GCTACCAGTC CATGACTAGC CTCCACACGA GCTCTGAGTC
26851 CATTGACCTC CCCCTCAGCC ATCATGGCTC CTTGTCTGGA CTGACCACAG
26901 GCACTCACGA GGTCCAGAGC CTGCTCATGA GAACGGGTAG TGTGAGATCT
26951 ACTCTCTCAG AAAGGTGAGC TTTCTTGGAG GCATTGATAA CATCTTCCCC
27001 CTCTTCCCTG CACTATGCCT AACCCCCACC CCATTAAATT CCCTTGATTT
27051 CACTGTGAGT GCCCCGGTGC AAAAAGATGT AAGACTGATG AAACCGGGCC
27101 TTTCATTTGC TCTCATTACC AAATTTACAG AGGAATAGAA TCATTAAAGG
27151 TAGGGTGAGT GGATAATTTT GTTAATATGA ATGCATACAT TTATACCCAG
27201 TAGGCAATGT GAATAAAATT CAAGGAATGT ATTTAGATAT TGAATGAGGT
27251 CTCCTGAAGA CATTTTAATG ATTTGGCTTA AGCTTCAGAA CAACACTAGC
27301 TCCTTATGAT GACTTAAGCA TTTTGAAAGA CCAAATTGAA ATTATTCTAT
27351 AGTTATGCTC AGAGCAATAT GTTAAATTTG TTCCATTTGT ACTTCTATGA
27401 AAAAATAGCA GATGGATTGC TGGGAAATCC TAGTTGGCCT GGTAAAAAA
27451 AAAAAAAAAA TCAATTGTCA GCCATGAATC ATTAGAGAAA ATTATAGTGT
27501 CAGTGCCATT TTCAATAGAC TGCTTAAAAA GTAATCATAT TACAAAGTGT
27551 TTCTCATTGG CTTTATATAT ATATATAAAC TTAAAGTAGA GGACATAGCA
27601 AGGCATTTCT TACCTAATAT GCTTACTGTG AAGCATCCCT TTTGAGCAAA
27651 ATCACTCTAA ATTTTCTCCT CAAAGTGATC CTCTCTTGAT TATACTGTAC
27701 TGACTCTTAC CACCAGGAAA ATGTCTTAAA ACCACTTCTT TTTCTGATA
27751 AATGCAATGC TATTTGTCTC TTGACATAAG TAAAGCTTTA AACATGGTCT
27801 TGGCCACATG TGGAAGAGAA TACTGGTCAC GTAAAATACC TGATATATCT
27851 TTCTATGTCT TCCCCGTCTT TTTTATTTT TTTTATTTT TTATTTTTTA
27901 ACTCTGATAT TGATGATGGC ATTTATTTTC TAGACCTTCA GCCTTACTCC
27951 CGGAATGATA TTTTAAACA TCAATTAAAG CCCTTAGCTA GACACTCTCT
28001 GCATTACGCC AGTTTCCCCT TAATGTAGGA TGTCCCAATT TGAAATTCCC
28051 CATTTTCTCT TGACTTTGTA AAATACAAA CCCAGAGCAA AACATTGCTT
28101 CTTTCCCTCT TTACTTCCTA CTTGCCTAAC AATGAGACAG GGACAGCCGT
28151 GCAAATGGGG CTTTCCGATG ATAAAGTAAT TTAAACACTA ACTAAAATAT
28201 TGGTGTTCCT TATGGTGGGC TGCTAATTAC AAAATACATT TTTCTCCTA
28251 AAGAAAAAAA CTGGGCCAAG GCAAACAGCT CAGTGATAGC AAATAAAATG
28301 TAACCATTTC CCTATGGTTT TGCTGTTATA TGCTATTATA GACAGCATAC

Fig. 2 (cont'd 15)

21/124

28351 GTAAAGACCA GTAAGGGTTC ATTTTCCAC CTAAAATGTC GGGCTTCCTG
28401 TAAAATCTTT GATTCTAGTT TCAGCACTTC TAAGGTAAAT GGGCATCTTC
28451 ACATGTCATT TATAAACTT CTAATGAATG AATTATATTA AAATAGATAA
28501 ACAACCTATA GTTTTAATGA ATGTATCCTA GATTGTATGC TCATATGTAA
28551 GGATTCTAAA TATCAACTTG ATAACCAAAC CAAACATAGT GCAAATAGGT
28601 TATCATTTAT TAACCACAAC CACCTTCCAC AAAACTGGTC ATTTTAAAT
28651 TATTAAGATA ATCTGCAACA AGTTGGCCAT TTAGCCATCA GCCTATTTCT
28701 TCAGCATTTA GACATTAATC CCAGATTCAG AAATAAAGTC AAGTAACTAT
28751 TTATAACCAA GTAACATTCA AATCAAACT AGATGAAAGA TTGGTTAGTT
28801 GCATAGCTAT AACCAAAATG CAGTTTAAAT ATTTTACTCT AATCTATATT
28851 TTAAGTGAAG TCAATAAAAT TTCACTATG GAAATACACT AGAAAATATG
28901 CAATTTCTTA TTCTTTTAA GCAGATTTAT TTATTGTACA TGTTCACTCT
28951 TTGAAATAGG CCAATTTTAT TTATGTTATG TTATGTTATT TATTTGTTTT
29001 GAAATGGAGC CTCACCTCTG CGCTCAGGCT GGAGGGCAGT GGTGCCATCT
29051 CAGCTCATTG CGTCCTCTGC TACCCGAGTT CAAGCAATTC TCATGCCTCA
29101 GCCACCTGAG TAGCTGGGGT TATAGGAGCG GACCACCATG CTGGGCTAAT
29151 TTTTGTATTT TTTGTAGAGA TGACGTTTCA CCATGTTGGC CAGGCTGGTC
29201 TCGAACTCCT GACTTCAAGC GATCTACCCT CTTGGCCTC CCAAAGTGTG
29251 GGGATTACAG GTGTGAGCCG TGGCACCAGC CTGAAATAGG CCAATTTTAA
29301 AAATGGGAGT ATTCCTACAT TAAAATGGCC AAATAAAGAC TTTTCTAAA
29351 ATAACTTTA AACTAATTTT GGATAAATAT GTTTGCCTT TGAGCCTTAA
29401 TAAAATGCAT TAATGAATAT TAAGCTGTAA AAAGTACATG TTAAGTACAT
29451 AGCTATAGTG TATAATATTA ATATTAATTA GTGCCTTCCA GTAAATTAAT
29501 AGATTAAAT AAATTTTAAAT ATAAGACACT GAGCTTTTTG TTTTCTTGAC
29551 AATAGAACTG CAAGCAATAG CAAATTGCTC TAATCCTTTC ACGTACATTT
29601 AAGAAAGTTT ATGACCTATT GAAGAGAAAA GTAGATCTAG TGGGTGATAC
29651 TGGCTTCATT ATGGTTAATT AATTGATCAG TAGAATGTCA GAAATGCTAA
29701 GAAAACCAA GAACTACACC AGAGAGAAAA TGTGTTAATG TAAATTTTAA
29751 GGCAAGTTAA TTAGCGATAT ATAATAAAGA TGTATATAAG TTCATGATTT
29801 ACCTGTTTGT CTACAATTTT AGATGATTTT TTGATACTCA TATTTAAATC
29851 GGTAGCTTTT CCTATAGATT TTAATTTTGT TTTAAATTCC TCTTCGTAA
29901 ATTAAATAAA ATAATAAAAT AACTTTTAA ACAGTTTCT CTCTGCAGC
29951 TGCTCTAGGT CATTGGTGGC CATTGAGCCA TAACTAGTCT ATATTTGTTT
30001 TGGGTTTTGT TTCATGTGTC TGACTCAACT AAATTTTAA ATAATTTGTA
30051 GTAACCAACT TTGCAAATTC TGGGTTTGT TTTAAATGTC AGATCTGGCA
30101 ACGCTGCCTT GACATTTCTG CCTAGAACT ATTGGCTCTA GGCAGTCAGT

Fig. 2 ((cont'd 16))

22/ 124

30151 GTCTGTCTGC TTCAGACTGT TGACTGAAAT CCCCATTCGT TTTCATGCCC
 30201 TATCTGGCCC TTGCTGGCAT ATGAGTTTGC AACCTTTGGT GATTTCGAGA
 30251 AATTGTCTAT GTTAGAAAAT CATTAAATATC TAGATTCAAA CATATTTCTA
 30301 AATAAAGCTT TAAATTATTA TGGTAACTTT AAATGTATTT ATTCTAATTT
 30351 TTTTCATTAA ATTGCTCTTC ATCATATAAA TATATAATTT TTATACAAC
 30401 GGATGAGTTT GGCAGAAGAA TACCAACTTT TCATATTCTT TGTGGCATT
 30451 AACTTTAACT TGTACACATG GAAATAAATA ATCCTTAAAA TGACTTATGA
 30501 CCACATAAAT GCCTTAGCAC ATGTGGTTCA TATTTGGAGA TTTCTCATAT
 30551 TTGTTCAATA TAATTTATTT TGTTTGTTTA TCCACAGTAC TTAAGAAAAC
 30601 TTCTATAGTC AACATATATA CTGTAACCTG CCTCTACACA GTATAAGCAA
 30651 TTACCTTACA TGGCTATTAC CGATAAAGTT AAAGTTGTAT AAAGCCTTTG
 30701 GATGCTTTTG ATTTTCAGTGC TAAATAATGG AGTACACATA GAAGAAAACA
 30751 TTTTAGCTTT GGTTTGAGTG ATCAAATTTT AGGTCAGCCT TTTTACATTC
 30801 ATGTTATATC ATCCCCATTA TCGGTATCCT GTGTATTTAA TTTTGATCAT
 30851 TTGATGTCCT AAAGGAAGAA AGCTATAATT CTGCAATTTT AATTAATTTT
 30901 ACACTTTGCT TATCCACATG CCAGAGATTA TAAAAGAAAT CCCTAAACTT
 30951 GTCCCACTTA GTTGTGATA TCCTCTTCCT GTATTTTATAG AGAGGCCATT
 31001 TCTTATTTTC TCTAGACATA GCTTTTCATT CCTTCTTGTT ACCAATTGTG
 31051 AATTCCTTAA AATAGAGATG ATAAAATTTA TAGCCTTTTA AATACCTAAT
 31101 TTATGATTTT TAAAAGATGG TATAGCTTAA TTTCATTAAA ATATTCAAAT
 31151 AAATGATACT AGAATCAATT AAGTTTAAAG CAAACATTCA TATATCTTTC
 31201 TTCACATGTG TAAATGGGAA ATAAACATGC CTTTTTATTA AAAATAATTT
 31251 GAAGACAAAA GATAAGTATT AAACAACGTT TTATACCATC TCTGTCAATT
 31301 GGAAGTTGTC ACTCTAACTT AGCCAGAGCA GATCTATCTC ATTTTGCATG
 31351 TGATATCATA GCAAAAGTCT AATCAGTTGC ATAGGGAAGG AAAAATAAG
 31401 ATAGTATTTA ATCAATAGGA TTCAGAGGAA AATTATGCTA ATGTGATTTA
 31451 ATCTATTTTC TAGTAATCCT ATCACTAAAC TGTCATTGAA TTGTACTGCA
 31501 TTAGAAAGGA ACTCAAATAT GTGTGACGGC AATGGACATC TTGTCACCTT
 31551 TAGTTGGCCT TTTTCAATGA GTTAAGCATT ATATGTGTGT TACCAAAAAA
 31601 TTATTTTATTA TAGTTCAGAG AACCATTTTT GTTGGATGTG TAATTTGGAA
 31651 GTTTTGTTTA CATTATGTCC TTAGGGGTTT TCTTTGTTTT AACAGCATGC
 31701 AGCTTGACAG AAATACACTA CCAAAAAGG GACTAAGGTA TATATTCCTC
 31751 TCAGCACAAT TGCTACCTCT CTGTTGTTAT GTAAACTTTG TGTGCTGTCT
 31801 CTCTTCCTTC TTTGTTTGTG TGCAATGTAG CACATGACAT TGAGGACGAA
 31851 ATCACTTTTA ATTTTGATGG TTTCTCTGGC CCGAACAGTT GGTGAGATAG

Fig. 2 (cont'd 17)

23/124

31901 CCCCTTAGGT AGAGATACTA GTAGAGATTG AGGCTGTCTC TCAAATTAAA
 31951 TAAATTCCAA TGTGAATATC ACTATTTTGA AGAAATAATA CTAAACAAAC
 32001 AAACAAACAA AACAAAAACA AACAAACAAA AAACCTGTCC CAGGCATTAC
 32051 TTTTTTGGGG GCAGCAACTT TGGTAGAATG CAGAACTCAC TTCAACAAAT
 32101 TAAAAATAAA TTAACCTCTC TAACTTTTGC CTATTAGAGT CATATGCATG
 32151 CAAATATTCA AAACCCATGC AGTCTACAGA TGTGGGCAGT TAATGTTGAT
 32201 AGGTGAAGG ATGCTACAAT CTGAATCAAA GAAAACATAT TTTTCATCATC
 32251 ACAGGACAAA TGCTGTAATT AAGGTGTGAT TTTTATAGAA TCCTTTTGAT
 32301 AAAATCTCAA AATTGTTTTA ATTTCTATTT TGCAGGGGTA CTGCTATCAG
 32351 ATCAATTTAA ATCTGAATTA ATCTAATATC ATTTAATAAT CTCAAAATAA
 32401 TTATTCCATC CATAATAAAA AATAAAATAA AAATTAACT TATGGCCATC
 32451 TTTTACTGTG TACTTTTATC TGAGGAAGAG ATAGAATGAT CTACTAATAG
 32501 AGGTATAACA CTGTATGTGT ATGAAAAGTT GGCTAATTTT GGTGCTAAGA
 32551 ATTTACTTAC AAAAAGAAAA AGAATATACT TAGTTTGGTG AAACACTGAA
 32601 TAATGGCGAA ACTAGGTCTT TCTCCATTAT TTTTTTCTC TCCAATTTTT
 32651 CAGCAATAGC AAATAGCTGG CAATTATTCC ATGTTAATAT TTTGATCCAG
 32701 AAATTTATGT TCCAGTAAAG CGAGCACATC TCCCTCCTTA TTTTGTAAAT
 32751 CTAGGCATGA TGTCAAGTGG CAGTTTAACA AAAGAACTGT TTTTCCTTTA
 32801 AAAAAAAAAA AAAAACAAAA GCTGCCAATA TGTATTCCAT TTCCCTATGC
 32851 CTTCGTGAC CATCCTTCAT TTCCCTTGGC CCTGGCCAC CACTGTCCTC
 32901 CATTTGTAGT CCATGTTTTT ACCCTCTTTA CATCCTTTCT TGCCCTGTGC
 32951 TTTTGAGTTC TCAATTAACT TGGCTGTCTG CTCATTGCTT ATGATTCCA
 33001 ACTGCATATC TGATAGAAGC ATAATTTTCT CCTCAAAACC CTTTATCTTA
 33051 TTTTTTTTCC CTATGTGATT CAAACAGATG GCGTAAGATC ATCTGGAAGA
 33101 ACTGAGCAAT TATAATTAGA TTCAATCTGT TTGAAATTGT TCATTCTGAA
 33151 TAGTAACCTC CTCTGAATTG TTTTCCTGTC CTGGCATTGC CTTGCCCTTG
 33201 TAGATGTGCT TAAGTGTCTAT AGCTGTGCTG TTTTGCAGAT ATACCCCATC
 33251 ATCTCGGCAG GCCAACCAAG AAGAGGGCAA AGAGTGGTTG CGTTCTCATT
 33301 CTACTGGAGG GCTTCAGGAC ACTGGCAACC AGTCACCTCT GGTTTCCCTT
 33351 TCTGCCATGT CATCTTCTGC AGCTGGAAAA TACCACCTTT CTAACCTGGG
 33401 TAAAATATTC TAAAATATTG ATTTTGTTTT GTTCTTTCA CCACCCACTC
 33451 TCACAGAAAC CCTGGAATCT CTCCATAACA CAACACGTTT TCATTTAAAG
 33501 GGAGGGATAA AAGCACTTTA ACAGTACCTT TCATTTGTGT CATTGTTTAC
 33551 TCTTCACAGA AAAATCTCCA AACATTATGC TATTTATTGC TCATGACAAA
 33601 TGETTAACAT AGATTAATAC TGTGGTTGTT TTCTAGTCTA GGCTCCAGAG
 33651 GCTCAGAAAG TTCACTTGAC TTGAAAAAGT CTTACCATTA CTAAGGGTTC

Fig. 2 (cont'd 18)

24/ 124

33701 AAGGCAGTAA CCAGTTCAGA ACATCTGACT TTAATCCCAG GGGCCTTTCC
33751 ATTCCATTTA AGAATCCTCT TAAAAACAG GAAGGCATCT CCTTATTTAT
33801 TTGTCTGAAA TATTAAAACA TCCTTAAAC AAAATTAGTA ATCTTTTGTA
33851 GAAATAGAA ACAATTAGGA AGAAAAAAT ATGTAATTCC ATGACTCAAA
33901 GTTAAC TTCT TTTAACACTG TTAAAGTTAA AACTCCTTAA AATTCATACA
33951 AGAATTTCTG TTAAGACAAT ACTCTGAACA TTTTCAAATA GATACAATGA
34001 AAAATAAATT ACCAACTTAG TCATTGGGTT ACTTTGTATT TAACATCATT
34051 TGTATGAAAT ATAAAATCAT TTGCATAAAA TTTCATTAAA AGCACTCTGA
34101 GTAACAAAAT AATTAAAGAA AACTAAACAT GCCAGATACC ATTTAATAGA
34151 TTCAATGACT TTAATAATAT ATTTATTTTC TATAAAGTCA CATATAAAGT
34201 ATTTTCATTA TTTTATGGT AAATATTTTT ATTATTAGTT TATCAGAAAA
34251 ACTTGTACAT AAAGATGAGT ATTGATACAT AATCTTATTA GAGCCAGAGA
34301 CGATCATTTCC TTCTAGAAAA ACACATCTCT GAATTTAGGA CGGAGGACAA
34351 TGAAACAAGA AATTTCACTT TATAATTTAC CTTTGTCAA CTATCCCAGA
34401 GCACATCAAT TCCATCATGA AAGTACTCTT TTGACATTAT ATAAAAAATT
34451 AGTAATAGAA AACACACAAT CCAAAACCTT ATATTTTCTA AACTTCAAGT
34501 TAATCATCAA CTTCTCTTAG ATTTTGAAG ACCTGAAAAT AAACATAATT
34551 TCAAATAACA GAACTCAAAC ACCATATACA TTTGTAATGA GGCACAACAG
34601 TCAATTTTGA GCCTTGTTAT TTCCAGGTTT TAGCTGAATA ATCTTCACTG
34651 CTTTCTTAGC TTTTGGCCAG TCTAGTTTGG GGACTATTTT GCCTTACTGG
34701 GCCTAAACAG AGTGTAATAT TAAAATATGT TAATAAGCCA TACTGAGAAT
34751 AAGATAAATG CAGGTTTCTA ACTCCTTAGG GACACAAGTG GGGACAACAC
34801 ATTCCATGAA CACAGGTGAA TGAATGCCCC TAGTTTCTCT GAGTTGGACA
34851 ATTTTCATGCG ATCATTTTTT TCTCTGAGGC CAAAGTCTCT GGTTTGATCT
34901 TCCTAGCAGC TTCCAGAACA GAAAGTGAGT TTACTTTGTC TCCATATTCT
34951 TTTTCTCCAT GCTCGGGAAT CCCCTGCTTT CCTGATCCCA CCACAAAAAC
35001 TCCCCTGAGG ATGAAGCCTT GGCTTTCCAG GCTTCCAGGG AAGCCTCGAT
35051 TCCTGGCTGG AGGTAGTTGT ACCACACTCC CAGAGGGCTA AATCCCATAA
35101 ACATCATCTT CTGTCTTTGT AGATCATAGA ACTTTTTATT ATCATCCAGG
35151 AAGATTTCTC TTTTGAAACA AGGCTGGAAA AACTTTATGT CAGTCCTGAC
35201 CTGCTCTTTA ATGACTGCGT AGAGGGAGAT GCCCAGCTTA TCCAACTTGG
35251 GTTGCAGAGA GGACAGATCT GCAGCCCCCTC TTGCCAGAGA AAACATCCTG
35301 GCACAGCCAC AATCACAACCT CCATTCTTCT CCCGATAGCT CCTTTGCTTT
35351 GAAACTCATT GGTTACTTCT CCAGTGTTTT CAGGTCTATA TTCTCCAGGT
35401 ACTCCAGCAC CTCTTTCCAG GGCTTGGACA AAAATACATC TGTGTTGGCC

Fig. 2 (cont'd 19)

25/124

35451 AGCATCAGTG CCAAGGCAGC AGCCTCCAAG GGCTCCTGCA CCCATGGACC
35501 ACATCCACAC AGAGAAGCAC CTTGGGTCCT CAAGTGCCTC CCTCTTCTTC
35551 CCTTCTCCCA AACCTGAAGC CCAGACACTA AGGGGTCAAA CCCTCCTGGG
35601 CCCTGAGGGT TCCAAGGGCC TCATTACTTT TTCTTTTTTT CACTGGAAAA
35651 AAAATTCTAA TCATGCACCT ACAGAAGATT GACATTTTTT AGTAAGTTGG
35701 ACTTTCCAGC TTTCAGCCAG GACAAGACTC AAGGCTATGT CTTTCTATT
35751 GCAACCCTTC CCACTATATT GAGTAGGGCT TTTAGCAATT GAAAACAATT
35801 ATTTTGGTCA TGGTTTCATA TAAGCTAATG ATTTTATATC AAACACCAAG
35851 TTTTGTTC CTAACCTATA TAGTGATAAG AGAATTTACC TATAATGCCA
35901 AAGAATGTAT AGCTTTTATT TGCTTTAAGA TGCAGTTGAT TTTTAAAAA
35951 AGCGAAAAGC CTAACACTTT AACTTCAAAA AATGAATTTA AAATGTTTGT
36001 GTAGGTCATA GGAATATGAA AAAATTTTAT ACAACATCTA AAACACACCC
36051 AAATCACCTA AAGTGCTATA AGCTTGCTAA GTACTTCATG TCTCCTATCA
36101 ATTCTTTCAT TAATTGACGT TAATTTGATT AGTTGACTCC TTCTTCTATT
36151 TTTCTCACC ATTATTATTC TGATTAAATC CACCTTCATT ATTCCTTAGG
36201 AACAAAAAGA CTCACCACTT AACTATGTCT GACATTGGTG AAGTCGTTTA
36251 AACTTAATTT TCTTATCTCT TGAATGGATA CATAATACCT AGGTATATT
36301 GTAAAGAATG ACGGATATAG TGTATGTAAA GATGGAGAAG TGTGTAAGAC
36351 TTGACAGATT CTGCCAAATC ATTATTTTCA CTGGAAAGCA TGTCTTACAC
36401 GATCATAGAG TAGCATTTCAT CAGATATGCC TGAGCTTTGT CTACATTTAA
36451 TTGAGTAGTA ATTCGCAACA CAGTAACCAC AGGATTTTAT GTAAAAGACA
36501 TTCACAGATT GTGTTTTTGA AAGATTGTAT TTTTGAAGTA CAAAACATG
36551 ACATTGTTAT CAAGGACTCA TTTACCACAA ATATCAAATA TTTGTGCAAA
36601 GATAAGTTTA TGCTAAGATT TGCATAAATT AAAGTTAACA TGGCAACTGA
36651 AGCTAACATG TCCATGGTCA CAATGTGTGA AAAAATGAAT GGTTCGTAG
36701 CACACTTGGG AATGTATTTT ATTACATAGT TTTCAGAGTT AAAACACAAT
36751 TAATAAATGA AATGTGAATT ATACTTTTAC TGACAACAAA GCTCTCTGTA
36801 GAGCTTTAAT GTTCTAATGA ATTAGAAAAC CACTGATCAA ATACATCCCT
36851 TACATTTTCAT TGCTATAGAA ACCAAGTCTG AAAGGTTAAG TTTACCTTTC
36901 TAGGATGTGG GTTTCCCCC TTAATCTATT GTGGTTTATA TCAGAGATCT
36951 CTCAGCTGTG TCAGACAGGC CATGACTTAA GTGACACTGC CCTCTTGATT
37001 CTCTTCATAC TTTTCCAAC ACAATTCTTT CTCCTGGGGT TGCTCATCTT
37051 AACATAGCTG TATCATTTAT TGTAAGACACA AGGTCACTTT TGAGAGTGAA
37101 TGGGACTATA TTAATAATTG TTCCAGGTAT TAGGTGCAAA CCCTGGGCAA
37151 TGCAATTCAT CCTCCATCTC CTCCTTATAT TTATGTGTTT ACCAAGTTGT
37201 TTTCTCTGTA GACTTTTTTT TATCCTAAAC CCTTTTCTA TGTTCTCATT

Fig. 2 (cont'd 20)

26/124

37251 CACAACCTTTA ATTCTAATCT CTCAAATCAA CATTTCACTT TCTGTCTGAG
37301 ACCTTTTTTCA GCTCTAAAAC TAAAATCCCA TCAGTGTGCT AGACCATATA
37351 GCCACCTGAA ATCAAAGTCT TTTCTTAAGT TCTTTCTTC TATTTGTCTT
37401 ATAATTTTCAT GTATCATCCT TCTCTCTACT CTAGCACAAA ATCTGTGTAA
37451 TCAATAGTCT TACTTGAAAC TGTGCTCTTC ATATTGTACA TTTTCAATAG
37501 ACAGGAACCT GTGATTTTAT CTTCAGAATA TCTCTACAT CTGTCTCTCA
37551 TTTTCAGGGA CATGTGCCTT GCTGAAGCTT TTTTAACTAT AGACAATTGC
37601 AGCAGATTTT AAACGTATCT TACTCTGTCTG ACTCCCTTAT GTTTCACAT
37651 TTTCACCCAT TGGGAAGGTAT AAAAGAAGAT ATTCCTGTCC GTGTCAACAT
37701 AATCTCATGT ACCTCTCCAG ATCTTAGAAA CACGTATGGC TTCAAATCAG
37751 GCATTTGGAG ATCTTTATGC TGTATGGTTT CAGAGTGGAA AAAATGATTG
37801 ATTCAAAAAC ATAATATTTA AAGAGTTTTT ATTGTATTTA CAGTTCACCT
37851 GAACCTCTGT TCATTGGGCA AGAAAATGAG TACTCTTAAA ATGCAATAAT
37901 AAATTAAAGT TACTTTATTA TTAAATTTTA AATATATATA TATATACTTA
37951 CCTTAAATAT GTCCTCTTGT TGTCTTTTAG CATCACCCAT TTTTGATTG
38001 ACCATTATCT TTTCTGAATA ATCAGTAAGA TACAGGATTA TTATTAATGT
38051 TCAAAAGTTG CAGTATTCAT GTTTTCTTTA TTCTTTCTAC CAATTAAAAT
38101 GTGTTAATAT ATAAAATTTT TAGAAATTTT ACTATAAAAA ATCACAACAT
38151 ATATTAGAAA ATTAAGATCA CTACAATATG TCATATTTAG TAGACTACTG
38201 TGAGCTACTG CCACAGTAAA CTATGGTTCG TGTGTCGTTT CCAGCATGCT
38251 AGCCCTAGTA GAAACCATTC CCATTCAAGA AAGACTAACA AAGTATAGCT
38301 TACATAAATC AAAAAGTCTT TGGATGAAAC TTCATTTGGG AAAATAACCC
38351 AATCGCTACC CTTCAATTTT TTATGAATGA AAAAATGGAA GAATAAAGGC
38401 CTCTAAGATC CATTCAAAGC CAGGAGACAC ACAAGAATTT CTAAATAGAA
38451 GAGAAACAGA AGAGGTCATA GTTCTTGTGA GCCATCTCAT AACCTGGTGA
38501 GACTCATTGT CATGCCTCCA TGCATGATAA CAATCGCTCA GATTCATTTT
38551 TCATCTTGCC ACAAGGGTTA CATGCAGGAA CATTAATGTC AACCTGTCAC
38601 TTCTAATATC CATCTAATAT TCTCTAAAT CGATGGATCC TTTTGATAT
38651 GGTGATTGTT AAACACCTTT GCATAGGAAC AGTTTCTATG CTTTGTACT
38701 CAAATCTTCC TCTACCTTGA ATCCTTTCCC ATCTTCGTGT TCAACCTTCA
38751 ATCTTCTCAG AATGAACTCC TGTCTTCTAT TCTTTCGGAA GCATAGAATC
38801 TCACGGTCAG AAGAGACCAC ATCTGGTTCA ACCCTTCATC TCTTATGTAA
38851 AATTTTATGA CATCTCTAGC TTCTTCTTTA AACCACCAA TGACAGAAAC
38901 TACTAAAATC TAGAAATAAC ACCTTTGAAA TTCTTTCTTT AAGAGATCAA
38951 ATAAAATTTT CCTGAATCTT CACCTATTGT TCCTAGTTAT ATATATCCAG

Fig. 2 (cont'd 21)

27/124

39001 ATTCTACAAA ATAAGTCAAA GTTAGATTGC ATATGACAGC TCTTCATATT
39051 TAAAACAATA TAATAAACTC ACTAGTTAAT GTCTAGCTGT AGATGCAAAA
39101 GTAGAGAGTG ACTTGGGGTT ATTTAAAAAC CCAGTCCAGC CAGACACATT
39151 GGATCATGCC TGTAATACCA GCAGCACTCA GGAGGCTGGG GCAAGAGGAT
39201 CCCTTGTTCCA GGAGTTACAG GCTACAGTGA GCTATGATCG TGGCACTGCA
39251 TACTCCAGCC TGGAAGACAG AGTGAGACCC TGTCTCACAA TAATAGTATT
39301 TAATAATATC ATAAAAACCC AGTCCACATT TATATAGGAT CCTGTTTTCC
39351 TCAAGTTACT ACAAATAAAT ATATAATCTT AATAAAAGGT TAGTGGCTTT
39401 GCCAAGATAG TGGCTTGGCT ATGCAAATGC AATTTAAGAC AAAGTTGGTA
39451 GCCCTCTTTT TCCTAATACA TTGCCATATC TGTTCCTCTT CTATTTGGAA
39501 ATTCTTGTTG GTCTCTTGGC TTCGAATGGA TCTTATAGTC CTTTTATTCT
39551 TCCATTTTTT AGTCATAAAA AAAGTGAAGG GTAGTGATTG GGTTATTTGC
39601 CCAAAGCAGA TGGAAAGCAA AACTACCACT AGAAGCTCTT TACCAATTG
39651 TGTTCCATTC AAAAAATTAT CTTTGTATGT CTTACATTTG TCTTCTACTG
39701 TATAGTTTTT CTTGTTCTAT TTACATATT AACTTTTCTC CTTCTTCAGA
39751 CATCTGCCCT ACTGGCTACT CTTGAAATCA GAGACTGTGT CATATTTTTC
39801 CTTCTATTCA ACTACAACAT CTAAAAGCAG ATCTGTCATA GTTATTAACT
39851 TAATTGAACA CTCTTAAATA GTTAGGTGTA ATTTCCAATG CAGAAGCTAT
39901 CAAAAGGGTT TGTAATGCA AACTATTCCC TTAAAATCT ATCCTAATCC
39951 TCATTAATGT TTCATCTTGA TAGAGCTAAG TATTATGTAT TGAAATTGTA
40001 GAAGTACACT TCACTTGGAT ATCTCTGCAA TCATTTAGGT AAGAATTATA
40051 CAAAGCCAAA AAGCAAATAA AATATCCTCC TAACCCTATA GATACGTATA
40101 CTAAAATGAT GCACCTGCAA ATTTGTTTAA TACTTCATTA ATTTAAACAA
40151 GAGTAAATTC ATACTGTGAA CCAAGAATAG GGTGACTTAC CCCAATCTTG
40201 CCACCTTAAA CATAAACATT TTAAGTCTTC AATGTCCTAC AGTGTACCTA
40251 CTGGCTGTTG TCACTAATCA GACCGAAATG GTACTAATGG TCACTGCAGG
40301 CTGAAGGAAT ATGCTTGAAA GATAGGCAGA TCCTCTCCCT CTCCCTTTTT
40351 TACTTTTTTC GCCTTTCCAT CCTTCTTCT TTTTTTCCAA TAGATTGTGC
40401 ACTTTGGAGA TTCATATTTT CTTCTTTTTC CATTACATTT TAAATATGTG
40451 ATTCTTAGTC CTATGCTTCC TTTTACTCCA ATCAATAACT GGCTCTATCA
40501 GAGGGTTGTT CTGTGTGTTA ATTCGGTTAA TACCAGGATT ATCAAGCACA
40551 GTGCCTTCCA AATGTGAGAT ACTTCTCTCC GGTACCTCT GGGTTTACTT
40601 TTCTGTTTT ACATTGTTTT GAGAGCCAGT ACTTGTTATTA AGAAGAAGTT
40651 TAGTGCCTGT GTCACAGAAA AAATCTTAGT AAATTTTGAA GTGATGTCAG
40701 AA~~C~~AACTCTA AGCCACTGAC GGATTCCACA GGGTTTTGAA AATACTCGTT
40751 AGTTCCTTT ATATCTTAAG AGGCTCCTGC CTGCTTTCTC ATATACCAGT

Fig. 2 (cont'd 22)

28/124

40801 AACAACTTG CTTTTCTTAA ATATGAGCAT TTAGAATATC TTTCTCAATT
40851 TTTCTGTTTT GCTTTTATTC CAAATTTTAC AACTATATTG TTTTCCAATG
40901 TAGTTGTACA TACAATCAAC CAAATCTTTC CTTAAATTGA TGACTIONAG
40951 GTGAGGACTC TTTGGCAATA AGCAATAAGA AAATAAATTG TTATTAAAAA
41001 TTACAGACTT AAGATACTTC TTTGGAAATA TAACATGTTT GTGACTTTTG
41051 ACCATCTCAT CATGATATGC TCATCTTAAA CAGAGTAGAA AATCATTTCA
41101 TATAATTAAC TTTATGGTGG GCTGCAGATA CCATGTATGT TACATTGTGT
41151 TTAGTTATAA AAATGTTTAT TATACACTAT TTCCTTATAA TCTAACTTTG
41201 ATAATAATGA TGGTCCTAAT CATGAACCTA CATCAATTAA GAGCTTGAAG
41251 TGACTGAGAG TATTTGCCTG GAAGCATTTA AAGCCCTTCT TGGGAAATTT
41301 AGATGTTTTA TATTTTACTT TCTTTTTGAT TTTGCTTTTT CCATTAAAGT
41351 GATTACTATT TTTAAAGAGA AAACCGAAAA CTCTAGAAAAG ACCATCTTTT
41401 CTTCATAACA GGTAGCAGAA AACACCATGT TATTACATTT CTAGCAAGAG
41451 CAGTAGAGGT GACTTGTTGG TTTTGTGTAC TGTGCTTTA GAAATTGATG
41501 TAAGGCTTCC CATAAACGTG CCAGAGGAAA AGAGGGACGC AATGGGATCT
41551 GTTATTGAAC ATTTCAAGAG CAGACTCTTA CCTTAAATAG GGACTCACTA
41601 TACATTCATG TTTTCATAAG TATTGGGATC ATGTTCTTAC TTTCTATCAA
41651 CCTGCTATTT TCATCTTTCA AGCTTAAGAG TAATAGGCTC TGTGTGTTTT
41701 GTTTTTTCAGT GAGCCCAACA AATTTGTCTC AATTTAACCT TCCCGGGCCC
41751 AGCATGATGC GCTCAAACAG CATCCCAGCC CAAGACTCTT CCTTCGATCT
41801 CTATGATGAC TCCCAGCTTT GTGGGAGTGC CACTTCTCTG GAGGAAAGAC
41851 CTCGTGCCAT CAGTCATTCG GGCTCATTCA GAGACAGCAT GGAAGAAGGT
41901 AAGCGTTGAG GGGGATTAAA GATGAAGTCA CTTTATTTAA ACCCTGAGAG
41951 GGAAACCATC GTGTCACTCA CATCACAAAG ATTCCTGAAG AGGAAAATAA
42001 ACTAGTGTA TATCATTTG GGAAACTAGA AGCTTGAAGA AGTTTATTC
42051 TGTATTATCT TCTATTTCTT TATGTATTTG GAAATATGCC AGAATTGT
42101 TATATTAATA CTTGGCTGTA GAAGAGTTA GACTAAATCT ACTTTTCCAA
42151 TACAGAAATA TACATATAAA CTATTTTCCC AGGTGCATCA AATATCAGAG
42201 CAAATGTTTT GTTTGACATT TTGGTTAAAG AGCCATAAAG ACACACAAAC
42251 CAGAAACATT ATTTTATGAA AATACCACAT GTTGCTGACT TTTATTCCCA
42301 GGAATCCCT CTGGTGCTAA TTTTTTATTA TATCATTTTA GAATTCATAT
42351 TGTACCTACT TTTTGCTTT ATAAGTCACT ATTTCTTCAT CCAATGGCAA
42401 TAAAATTGTC ACCTAACCTA ATAAATATCT TTATAGTTAT ATAGTTCTAT
42451 GTAAATACTC CAAATAAATC AGCTTGAAAA CCTCAGGAAG CTGAGTTGAT
42501 GCTCAAATAT ATATATTTTT GTAACTGTA GAAGCTCAA TGTCAAATTT

Fig. 2 (cont'd 23)

29/124

42551 AACAAATAATT TGAGAGACTT TTCTCTTTGA TTTAATGAAT TTTTTTAGTA
42601 TCCATAAAGA AAACCTTACAG CATACATATT ATAAAGCATG TCAGCTAAGG
42651 ATAAAATAAA ACTAGACATA CAAATTCAAA CTGATTAGAA TGAAATTATT
42701 AACCCTAATA ATTATGTTTA AAAGAAAAGT CTCCAAATCT TGAGACATAC
42751 CAGAGTTTAA GTCTTCAGCC ATCCATTTAC TTGTGGTATA AACTTAGGCA
42801 AGTTTCTTAA CCTTCTTATC CCTAAGTTCT GCATCTGTAA CTTCTTAGGT
42851 TTGTCACAAG GATGAAATAT GAGAACAAAG AATAATTCTG TTCCATGATC
42901 TTTTCCCTTC CTACCTTCTT ATTTAAAGTA TCTTCTGACT GAGGGGTTAG
42951 GCAGCAATGA AAATTGACTC ATGTTTTTCA GGTCACCACT ATGGATTCAA
43001 TATACTGGCA TTAAATCAGT AGAGAATAGT TGTCAATGCC TTTTGCAATA
43051 TTAACCAAAC CACTCAGTTC ACTGTGACAG ACAGTGAATT ATATCCAATG
43101 ACTCCACTGA TTTTTTCCAT GTAGATAGAC AAAATATAAC TACTCTCAAA
43151 TGTAAGGACC CTGCTTTCTG AAATGGTTCT GTTGCTCTCT TCACAGATAG
43201 GCTTCTTATA ATACTTTTAA AATAATTTGC TAAGCATACA GATGGCTTTC
43251 TAGAGTGTGG CATTGACAAA TAAAGTGATT TTTATATACT GGGAAATTCT
43301 GGCCTTCAAT GTATCAGGAT TAAATAATCT GAATTTCTGA AAGCTAGCCT
43351 AAGTGGGCAA GATGGCTTTT TTGTGCTCAC GCATTGAATA CTGAACATT
43401 CTAGTTCTTA AATGGCGATC TAGATTCAAG ACTTATTGAA CTAGATTGAA
43451 GGGACTTTAT TGATATCCTA CCTAATGCTC ACACTGACAG ATGAAGAGAC
43501 TGAGCCACAT GTTCTAAGGT CATAAACAGA AAGAATGAGA ATGAGATGGT
43551 CTAATTAATT GTCCACCTTT CCTATGGTAC ATCAGGGTAA CACTTTAGTT
43601 TACGAGGGTA TTATTAGAGA TAGAAAGAAT TTTTTTTTAA ATAATTGACT
43651 CAAATACCAA CATTTTGCAC ATTACATAGA GTAATAGCTT TGCCCAAGTT
43701 AGAAAACCTGG GGGTTCTTCT TTATTCCTCT TTTGACCACA TCTATATACT
43751 CAGTTTTAAA AAGGTTCTTC CTGGTATCCT TCAATTCCAT CCCCATGTTT
43801 TCATCTACAA GCCTAGTGCA GCTATTCCAG CCGTCTCCTG ATCAGGTCTT
43851 AAGCACCTCC CATATGTCCT TGTAGTACCC ACCATATTGA TCTCAGTAGC
43901 AATCACAGTA CTCTATTGTA AATATCTTTT AAATTATTAT CTTCTCTTTG
43951 AGCTTTTGGG ATTTTATCTT ATTTATTTT GTAGTTCCAG GATCTAGCAA
44001 CAGCTTGTC AATCGTTCAT ACTCAACTAA TGTTTGTTTA ATGCACAATG
44051 AGCAGAAATA AACATACTAC TCCATAGTAA AAAGAGGATG AACTTTTCTG
44101 CAAATATTAA TCAGACCAT TTTATCCACC TTTTGGGTTT AGTACATTGG
44151 AAGTATAGGA GTATAAAGCA GAATGTCCAA TGTTTACAGT GATATTTTGA
44201 AATAGATAAA AGCCAGTGCG ACATTTCCAT TCTCAATTC TCTGAGACAT
44251 CAQCTTGAAA AAAAAAAGTA TTTTCTCTT CCTAAAATTA GTAAAGGAAC
44301 AGTAATCCA CATTTATAAG AGTATGATCA ACGCATCACA GATAATGTTG

Fig. 2 (cont'd 24)

30/124

44351 TAATAACACA TTAGATAAAA GTGCTTATTT TCCTGAAATT ATATGGAGAA
44401 AAAAATCTGA AAGTGGACCT TTGTTGGATA CAAATGAAAT AAATAAGGTA
44451 CATACATTTT TTAAGGTTCG AAAGTTTATG GCAACTTTAG TTTGGGTTTC
44501 CATGCTATTC TATTTATTAT ATGGGAATTT ACTGTAGCTT TCAACATGTA
44551 CGAAACAGGC TGGTAGGGCT CATGCTTGTA GGCTTCTGTC TAATAACTTG
44601 GCAACTGAGG TACTTTAGGG AGTATGGATG GGGCTCTTCC ATGTCTCAAC
44651 GTCCTGACTG CCAAAAAATT ATAGCAGGCT GGTTCCTCAGA ATCTTATAGT
44701 TAGTTGTTAT TACTTAATTT CCCTAACCAC CCGTCTTTA CTTTTTCTGT
44751 AAAGGCTGGA ATTTTTGAGT AGACCTTATT GTTTTAACTC TATTGTTCTG
44801 TTTGTTTCT CCAGTTCATG GCTCTTCATT ATCACTGGTG TCCAGCACTT
44851 CTTCTCTTTA CTCTACAGTA AGTAATGGCT GTTAAGAAAA AGCTTGTGCT
44901 TTTGCCATGC ACACAGATGA TGAAATAGAT CATTTTACTG TGAACAGATC
44951 ACATTCATCT ATGACTTGCA CAGGAGTTGT GTAGCAAAAT AACGGCATAAC
45001 TCTAAGCTGC CCAATACCCA ATAAAGTGCC AGGTGCTCCA CCTGCCATTC
45051 TTTGGTCACT TACATGTGCT TTCACTTGGC TTTGTGCAC TCATCATAAT
45101 CAATGAGTGG ATGTAGAATT CGATTTCTA AAACCTACTG AGGTATGACT
45151 TGGAGTCTCT GAAACCATGT ATGTAGTCTG CTATACTATC ATTTTAGTAA
45201 TGACGAGTTG TCCATGTTTT GTTCTTTGAG CCGTGACTGT TAATTGTTCT
45251 ATAGTATTTT CTTCTCATTT TTTATTTTTA AGTTTATTGT TGAGAGGATT
45301 ATCGAAGGGT AAAAGCAGTA AGGGTAAAGG GTAAAAGCAT AAAAGAACCA
45351 GAGATGTTTT TTTTAAATA TACCTTTTGA AAGAGTGTGA TTTTTTTAAC
45401 TTTTATTTTT ATTTATTTTT ATTTATTAT TTATTTATTT TTGAGTCGAG
45451 GTCTTGCTTT GTCACCCAGG CTGGAGTACA ATGACACAAT CATAGCTCAC
45501 TGCAACCTTG AACTCCTGGG CTCAAGTTAT CCTTCTGCCT CAGCCTGTCA
45551 AGCAGCTAGG ACTACAGGCA CGCACCACCA TGCCCAGCTA ATTTTTAAAT
45601 TGTTTTAGAG ACAAGGTCAT TGCTATATTG ACCAGACTGA TCAATACCCA
45651 TGGCTTCAAG CAATTCCTCC TGCTTTAGCC TCCCCAAGTG CTGGGATTAC
45701 AGGTGTAAGC CAGCACACTT AGATAGAAAC TTTATTTATT AAGAGAAAAA
45751 TACCAGTGTT TCAAGTTCTT TTGCAAACGT GTGACATTAT AATTCATTTT
45801 TGACAAGGAG AGTTTTCTG TTTGGTAAAT ACAATTCTAT CTTTTTTAAA
45851 AAAGTAGCCT ACAGGAAGTT ATATTTTATG AGTGAGTCTT TTTAGAGCTA
45901 GGTTAACAGT GAGGTATATT TAAAAGCAGC CTAAGTGAATC TCAATGGGAC
45951 TTGAGTACTA TGAATAAGCC TTAATCCTGT ACTGTAAGGT TCATGAAGAG
46001 TTCATAGCCT CTGCTGTCAC TGATCAACTG AGCATCATGG GCAGTATTTT
46051 TTCACTCAT TATCATTAGG TTCAAATGTT TGTTTGAACC TTCTCTTTAT

Fig. 2 (cont'd 25)

31/124

46101 AGATTAATCT CATATATTTA CTGCCTTACA TAGTCATTCA AAATCTGACT
46151 GTTATTGGCA GAAGTAATAT TTTTCTAATC TCTCCTTTCA ATGATTAAAA
46201 TTACCCATAG CTTCTAGAAA TTAAGAAATC ACGATTAGTT TTTAGGTAAA
46251 TGTACTTTTT GTGCAAATGG ATAAAGTGAG GAATGTGTAA ACACACATGA
46301 AAAAAACACA TAAAAGAAAT ATATTAAGAC TTAGTGTTCC TCCTGTTGGG
46351 CCAGCACTGC CATTTGTTGG GGAATTGTAT TCTGATTAA ACCATTGCCA
46401 TTTACATCTA TGTGTAACAT CAAAAGATGT AGCATCATT TTTATTCTAA
46451 TACATACAAT AATTAATATT TGGATAAAGC TACCTTCATG AAACCTAAGA
46501 AAAACTAAAT TAAAAAGAAA GAAAGAAAGA AAAATACACT TAGATAGAAG
46551 AAATAAGGTC TAGTGATTGG TAGCACAATA GAGTGACTAT AGTTAACAAT
46601 AATTTATTGT ACATTTCAAA ATAGCTAGAA AAGAAGATTT GGAATGTTCC
46651 TAACAGGAAG AAATGATATT CTTCTTAAAT GAAGAATGGG ATATTCCACT
46701 TTCCAGATT TGATCGTTAC ACAGCATATG TTTGTATAAT ACCACATGCA
46751 CCCCATAAAT ACATACAACCT ATTGTGTATC CCAATATTAA AGATTTTTTT
46801 GAAAAATTTA TTCCTCAAGA AAAGGATCAT GAGTTTAAGA AAAACAGAT
46851 TACTAGTCTA CCAGTGTTCA GTAGACCTTT CTGTGTTAAT AAAAGTGTTT
46901 TGTATCTACA CTATCTAATA TAGTAACTAT GAACCATATG TTGCCATTGA
46951 TTATTTGAAG TATATCTGGC AAAGAGATGA ATTGACTTTT TTATTTTAAT
47001 TAATTTACAT TGAAATAGCC ACATGTGCCT AGCAGCTACT AGATTGGATA
47051 GTGCAAGTTT ATAGAGAACA CAAGGGGTAC ATTTGTAGAT AGGAGTGGA
47101 TGTCAAAATG ATGAGGATAA TTAGAAAGCA TACATGAGAA ATATTGTTTT
47151 AAGAGTAGAA TATGAAATGG GAACACAGAT TAAAATAGAG TATGTATATA
47201 TATACATATA TATGTGTATA TATATACATA TGTATGTGTA TATATATACA
47251 TATATATGTG TGTGTGTATA TATATATATT TATAGGCCAA TATATGGAGG
47301 TAGGGTATAT CCTAGTGTTA AGTGAGTAAA GAATGGATTA GGTGATCGAG
47351 CCACATGAGA AGGTGATATT ATTAGAAAAT TGAAAGTTGT ATTTGAGATG
47401 ATGAAAATGA TATATTTGAA TTGAAAAGTA AACTGTAGTA AAATAATTCA
47451 AATAAATGAA TATTTGGGGA ACTACTTAAG AGAAAAATCA TAAAACATGA
47501 GGAGTCATTC TTTCCCCAGT CCGCCATGAT CAGGCCTTAG GATTTAATTG
47551 GCAATGAGAA AATACCTATG AAAATGCTTT TTAACTATC ACATGAAAAA
47601 GCAATTTATT ATTTTTCATG CCTTCTTAAT AACTCTCAAT AGAGATTTAG
47651 TTGATTTGCA TTTTGCCTG GTTCAATCAA GAAATTATCG CGTGACATCA
47701 GGCAAGTTGC CAAATTTCTT TGGACTATAC CTATAAAATA AAATTTGAAA
47751 ATATTAGCTA GATCTAACCC ATTTGTCTCC GGATGTCTGC AAAGTGTTG
47801 GAAATCACAA GCCTAACCTG ATCTGCAGAG GTGTTACCTT TGGCAAACCT
47851 ATGGTTTTTG TGTTTTTTTT GAAATCTAAG GCCAAGCGCG GTGGCTCATG

Fig. 2 (cont'd 26)

32/124

47901 CCGGTAATCT CAACACTTTG GGAGGCTGAG GCGGGTGGAT CACTTGAGGT
47951 CAGGAGTTCG AGACCAGCCT GGCCAACATG GCAAAACCCC GTCTCTACTA
48001 AAAATACAGA AATTAGCCCG GTGTAGTGGC ATACGTCTGT AATCCCAGCT
48051 ATTTGGGAGG CTGAGGCAGG AGAATCGCCT GAACCTGGGA GGCTGAGGCT
48101 GCTGCAGTGA GCGCCACTGC ACTCCAGCCT GGGCGACAAA GCCAAACACT
48151 GTCTCAGAAA AAAAAAAAAA AAAAGGAAAA GAGGGAGAGG GGAGGGAGAG
48201 GGAGAGGGAA TCTAAGCCAA CACTGTGAAA TATTGTGAAA TATGGAGCTT
48251 CTACCTAAAA ATTCAAAATT TTAAATTCCT TTTAAAAATA ATTGGAATAT
48301 CTATGGAATA TCTAGCAATA CTAAGATGAA ATTCCTCTGG GTTTTCAGTC
48351 ACCTGTAATT GACACCTTTA GATGTTGGCA TGGGCTCTCA GGAAGCCACA
48401 GCCTCCACCA ATGCTTTTCT TCCTGACACT GAAGCTAAAT TTGGGTGGCT
48451 AGTTTTTCAT GTGCTGTTGC TTTCTCATG GGAAAGAAAT ACCCTTTGCT
48501 ATTTATATTG CTGTCAAATG GGAAATGAA AGACAGCCAA GGAAGATCAT
48551 GTGACTATTT AAATACTTCA AGTCCATTTA TTCTTTATTA GCCTTGTCTT
48601 GTTAGGCATT TAAATTTTTG ATCCCTGCAA TAGATGTTTT TTGATTAACT
48651 GTATATTAAA AACTATATTT AACCTGTTTT GAATTTGAAT TCTAAATTGT
48701 ATTTTTTCAT GAGAGCAAGT GTCATTTTTG ATTCATTGTG GATTGTTTAA
48751 CATGTTGCCCT AACAAATAGC TAATACTAAC GTCATAACTT TTTAATTAGT
48801 AAATTTGAAT GGATAAATGG CCACTTATTG GCTTATAGAA TAAATAAAAA
48851 CATTTTTATT CAGTCAAGTG TTTCATATTT TTTATCATCT CCAGGACATT
48901 GGGCTTGCTC AAAACCATTG TTAIAAAAAA AATGGCAAAT AATCCAGTTC
48951 CATCATGATA TCATTAATCC CACACCTAAG CTAAGTAAAA AAATATATTA
49001 ATATTCCTGC TCATTGCTTT ATTTTTATGG TAACACCCAC CTGGTATTAA
49051 TAACCACAGA GTACGAAAGA AGGCAAAGGT TAAAGCAAAT AATAGTTTTG
49101 AAAAATTGGT AGTGAAAAAA GTCATGCTAT ACGGTATGTA TATAATAGAT
49151 ATTTAATGAT TATGCTTGCT ACTAGTATAT GTAACAGGAC TATTATAGAT
49201 TAACAAAAAT GCGGTGAGTA TATTTCTTGA TTATTTTTTA AAAGAATAAA
49251 TTATTATTTA AAAATACATG AATTATTTAT TGATTCTTGA ATCTTTACCA
49301 GCTTTCTATA ATTCTAGGAA GCCTAGAAGC AGAATTGGGC AGGATAAACT
49351 GGCAAAAAAT GTAAAAAGTA GGCCGGGCAC GGTGGGCTAC AGTGAGTCGT
49401 GAATGCGCAG TGCACCTGAG TGATAGATCA AGATCCTGTC TCAIAAAAAA
49451 AAAAAAAAAA AAAAGAAAGA AAGAAAGAAA AACAACAACA AAAACAAAAG
49501 CAAAGTACTA GGGAAACTA ATAGACATAG TTACATAGTT AATTGTGCCA
49551 TATGTTTTAA GGCAATGAAA CTTTTATCTT AATATTCCTT GCTTACTTTT
49601 TATTCAAAAA CCAAACTGTG TATAAACCT TAAAATTATT AGGATCTAAA

Fig. 2 (cont'd 27)

34/124

51451 AAGAATACTA TTTTAATTTTC CCTTGGTAAA ACATTTGTCC TGGTCAAAGA
51501 GAGCAGGAGG ACTTTAATTA TGACTTTATT CAAGGTGAGG TAATGGCTGT
51551 TTGATTGGTT TACTACTGAGG CAATCAGACA ACAGAGAAAA AAAATGCCTT
51601 AACAACAGCT TTTGCAAAAG TATTCCTTTC CTTTGAAGTC TTATTTTATT
51651 AGCCTTTAAA AATAAAATTT GTGCTATGTT TAAAAATATT TGAAAATTAT
51701 TGATTAAACC AATTTGTCTT TATAATCTCT GAACCAAAGA GTGGATATGA
51751 TTTTAAAAA TCAAAGTGGT TTTATTTACA TCACATGGAC ATGACAAAGC
51801 TTCTAACACT GATCATAGTA TAGCTACTGA AGCATCGAAA TGCTACATCT
51851 ATTTGCCTTA GTAGTAGTTA TTCAACTCCC CTTTTATCAT TGATGCTGTA
51901 TCATGAGTTA TGGTTTAAAA AAACAATTTTC AATCACTTTA CAGTTTCCTG
51951 GATTATATTT TAAAGATACT GGAATCATGT AATAGAGACT ATTTAATTTG
52001 AGAAATGCTC TTTGAGTTTG GATTCATTTA TGAATAAAAT AGACGCTGTA
52051 TTTTCTGAAA TCATTCATAG TCATTATCTT ATAAATGTAA AGCAAATGTT
52101 ATTTTAGACT GGGGTGTATC TGTCCGGAA AAAAAAAAAA ACAGGAACGA
52151 AGTAGAATCA CATTTGGTGA AATTATATAA GTGTCTACTG TTTCCAGCTT
52201 AGAGTTCTCT ACTTTGTTAG AGTGTGTGAG TTGACCACCA TTTATTTTCA
52251 ACAAATCTA ATGCCCCGGG CAAAACTAG ACAGTTAATA AACTATGTCA
52301 AGAATTCTCT TTCAAACGTA GACAGCATT CAAAAGTTCA ACTACAACTA
52351 TAGATAAGAT TTGTTTTTGA AGAAATGAGA AGCATCAAAA GTAGAATGTT
52401 TAACATCCAA GTAACGAAA TCCCTTGAGA CTAGATATAT ACTTATAGAA
52451 CCTAGTGTCA GATTGTTATA AATGTTCTAT CCTTATTAGT CACAACATGA
52501 GACTTGCAGA ACAAACGTA GAAAGTGCTT GAATTAAAAC TTAAACATG
52551 ATATAATATA TCCTTACCCT TTTCTGTTTC AGTTTTATTG GAGTGTGAAC
52601 TTAACATAAA AGAAAGATAC CTTAGAAATAT ACATTATATT GGTTTATCTA
52651 ATTAGTTGCA CCTATCATG GTTTTTTCCC CTGATTTTTA AGATGTGGAT
52701 AAGCTATAAA GCATCTCTGA GCTAATAATA ACTCACTAAA TAAAGGTCTT
52751 GATAATACAG ATTTGGGAAG GCTTCTCTGC AGTCATTGAA ACTCCAGCCA
52801 ATAACAATTT AAATGTGAAC TGATTAAATG TTGAATTAAG CCCAAGTTTT
52851 AGTGATTGCA GGATATTCCA TAGCCTTTGA GAAGTTTTCA AACTATGAGA
52901 AATTAAATG TACAGAGGAA AAAAAACCT AAGATTTTCT GAAAAAGAAC
52951 ATGGAGTATC TTTTACTAAA AAAGAACAAG AAAAATATGT GTGTATATAC
53001 AGTTTTTATA AAGAAAATAT TTTTCTACAG TTTTATTACC ACAGTTTTTC
53051 TAGAAGGAGA AGAATCAATA CAGAGGGTAA ACTGCTCTTG AGTCATTGTC
53101 CATTTGAGGG ATGGCAAATG GAGCAAGTGA GCGTACTTTG ATTTGTAGAT
53151 TAGAGTTTGA CACATAACAC TTTGCTTTTG AATGACATTT GCTTGTTACT

Fig. 2 (cont'd 29)

35/124

53201 GTGGAGTCAG TGTTCATATC CTTTATTTTC AGGAGTTGCT GCTGATACAA
53251 TGGGGTTAGA ATGAGCTAAA TACAGCATTT GCTTCTTG G TTTGAATTCT
53301 GGGTTTTAAG TAAAAATCTA CTTGCCTATT CCATTGATTT TTTTAATTGC
53351 ATTCAGCAAA TCCATAAACT GCGGAGAGAG CTGGTTGCAT CACAAGAAAA
53401 AGTTGCTACC CTCACATCTC AGCTTTCAGC AAATGTAAGT CACTTCATTT
53451 TTAAAATATA TTACAACAAA TTTTATAGA GGAAAATGAA ATCATTTTAG
53501 TAACAACTT ACAAATTTTC AGTGCCTGAT ACAGACTTAG ATTACCAACT
53551 AGCAGGACTC ATAAAAAGTT AACATTTTTT GCCTACTCAG TAATAAAATG
53601 TAAATCCAAA CTGATGAGAG GCAGCAATAT GGTAAATG GCTTGTGTT
53651 TCTAATAAGA TTGGAAACAA TAGTAACAGC CATATGGGTT ACTTCTTTTC
53701 TTGTTTGCTA TTTTATTAC TCCTCTTGCA TAAGATTCCC TGACAATGTA
53751 AGAGGGGTTG TTAGTGTGTTG ACTTTGGAAG ATAAAATATT CCTGTGCCCCA
53801 GCCTCCTTCA TCTCAATGTA TTGAACAATT TGTTAAGCAT CCAGTTAATT
53851 CTAAAATATG AAATTAGGTC TAAATAGGGA TAGCTTAGCT GCACTGTGGA
53901 TGAGATATGG TTTGCTCAAA AAACCTTGGC AGCCTTCTCA TAGCAATTTA
53951 AAAGGGTACA CTTTTACTGG CACCAGAGCA GCCCAGGATG GCAGAAATGA
54001 TGACAATGAA GACCGTCAAT TAAATTAACA TTTACTGAAT ATCTTCCACT
54051 GTGTCAGGGA GCACTCAGAG TAGATGCAGA ATGATAAAGG AGAAATGTGG
54101 CACTGTTCCT AGTCCTGAGG AGCAATGGTG TTAAGAACAG CAGTGAGGGG
54151 TAAGGAAATG CCTGCTATTT TGCCATATGT CTTACCTCTC TCACTCAACA
54201 GTCCTTTGCT CAGTCTGCT GCATAGCTTT GGGCTGCTC TGTGCCTCCC
54251 CACCCCTCCC ACTGCTCCTC TACTGAGTTT TTCTATCTCC TAGACAAAGC
54301 ATGATATGTC AAGAGTGAGC AGGTGCAGAC CCACAGTGTA AGACTTGAAT
54351 AAGAGCCATT TTTAAATTTT TTTTAAGCTA TCATTGTGCA ATATAAATTC
54401 TAAGTATGTG TATCATTTCA TTCACAATGT ATTCATTTTA GCACTGTATT
54451 TGAATTGATT TTATTTTCTG AAATTTGGGA GAATTAATTT TGGATTTATT
54501 CTATTTATTT TTAATAGATG GTGTTAGGAG ATTCCTGAAA ATAATAGCAG
54551 TTTTLAGATA ATTGTTTAAG CAATATGAGA AAATAAGGGT ATTATTTAAC
54601 CTGTTGTGT TTTTAAAGAG ATAGTCCAGA GGCAACCGTA AATTTTATAA
54651 TATAGGCTAC ATGTATAGAA GTATGAAATA TTGTTGTCTA GGTTCCTGAA
54701 TTTGTACCCA GAGGAAGTAG AATAATGTAA ATGTCAGAAC CTCCTGGGTT
54751 GTGTTTATCT GCAATAAGAA AGGCTCAATG GCAAACCTTA TTTATTAGAT
54801 TGTCAGGATA CTTGCAGATG TCTTGAATGA TTA CTAGGG TTTCATTTTA
54851 TTTTAAATGT CCCTTGGTTG AGCTCATCAT ATAATTCAGA TATTGGAATA
54901 ATAAATGGCT GCTAGACATA GTGGAAGATG GGCTGATACT TTCCATTTGA
54951 AATGTAATGA TGCTTATTGT CTTCAAAAGA AAAAATAAA ATGGTATTTT

Fig. 2 (cont'd 30)

36/ 124

55001 ACATTTTTTT GTTTTTGTTT TTGTTTTTTT TTCTCTGAGA ATCTCATTCT
55051 TACTCATGAT TATTGGTTTC TTGTGTACCA TTTCAACATT TTTCTATTAT
55101 ATGCTAATGT GTATATATAC TTAATACACA CGTGCAAAAG CTTCCACACA
55151 CACACACACA CACACACACA CACACACACA CACACATACA CACACATACG
55201 GAACCAAATT CTAACATAGG GGAATAATCT TCGGAGTGAA CTCTGTGCTG
55251 CTGTTTGAAA ATGGAGATAT AATTTTAGAA AGGTTCCTGC AGTTGGCTAC
55301 CCACCTCGTC TGCTCTAATT ATGCTTGTC CACTATTTTC ACTGATGTGT
55351 TTTCATGACT TTAGGGCATG AATTCTCAGC TGGGTGTTAA TATGACCAAC
55401 AAAGGGTGAA AACAGGTTCT TGCATTTTTT TAAGTACTCT TTTTATGTGA
55451 AAAGCACAGA TATGCAGATA ATACATAACT GAACATCCAG CATATCTGTG
55501 GCTTTAAAAAT ATCACGAAGA AGAGCACAAT TAGGGAAAAG AAAACATCTA
55551 TAGTGTTCCT CTAGGGGAAC AATCATTTAA AAAAAATAA AAATAAGGAA
55601 CACAGACTAG AAGCAGCAGT GCCAAATAGA TAATTCATGC TAGTCTTTGT
55651 GTTAATTTAA AAAGTGCTAG TCTTGAGAC AAACGCCCAA ATTGCTCTAG
55701 GTTCCACTCA GCTGTATGTG TTATCATTAG TATTAACCTT TGCACGCTGA
55751 TGGGAGACTG ATATATATCC TGTTTTATGT TCCTTTAAAC AATTTATAAT
55801 GTAATTTAGA AACCTTCTCA AATCACATTA GATCCACACA AAAACCTGTA
55851 CATAGCAGCT TTATTTTTTA ATAGCCAAAG AAAGGAAACA ACCAAAAATA
55901 TCCCTTAATA GGCCAGTTAA TAAACAAATT CTGATACATC TATATCATGG
55951 ACTACTACTC AGCAATATAA AGAAATGACT ATTGATACGT GCATCAACTT
56001 GGGTGGATCC CAGGGGTATT ATGCTGAGTG AAAAAAGACA GTTATAGAAG
56051 GTCAAATTTT GTATAATTCC ATTTATATAA CATTCCAGAA ATGGCAAAAT
56101 TAAAGAAACA GAGAACAGAT TAGTGATTGC TAAGGGCTAA GGATGAAGGA
56151 GAGAGAGAGG TAGTGTGACT ATAGGAAGAG GGAGATCTTT AGTTTTGTAT
56201 TTTGAATGAG ATGGCCATCA CATGAATCCA CATATGTCAA TCTATTAATG
56251 TAAATCAATA TTGTATTCCT GGCTTTGATA TATAATATAA TTTTATAAGA
56301 TATATAATCA TTGGGGGAAA CTGGATGAAG GATACAAGGG ACCTCCCTGT
56351 ACTATCTTTG CAACTTCTTG TGTATATAAT TATAAAATAT ATAATGTATT
56401 AAAATGTATA AAATAATATT TTAAGTATCA GATACTGATC TTTACTCAGT
56451 ATATGAAGTG TTCTATCATA ACGTAACATG CTTTTCCTTT ATTTGTGGTA
56501 TTTTAGTTTC AAATAAAAT ATAAATCACC TAAAGATCTA CGACAGTTCT
56551 TTTGAAAAAA AATCTTGCTT TTAATTTCCC AGGAGTTTCA ACCTTAATCC
56601 TCTCTTTAGT GTTCTTTTAT TTGGTAGTGA TAGGGACTAT CAAAGCTTCT
56651 TACCATCAAA TACATTTACT GACTAAAAAT AGAAAAATAA TTTACATTGT
56701 AAAAATGTAC AAATTGAATG ACAGTCAAAA GGTACAGGTA ATGAAGATAT

Fig. 2 (cont'd 31)

37/124

56751 GCATTAACAT CTACTTTTAA AAAAAAGTTT ATTAAAATTC TCTTTTAGAC
56801 TAATGCAGTA TCTGGGAATT TATATAAATA GATATGTATA TAAATGACTA
56851 TTAAACAATT TTAATGTCAG TTATATTTTA AACATTTTAA TAATATTGTT
56901 ATAACATATGG GGGTAAAATT TTGTATATAT CTGAACATTT TTGTCTTAA
56951 GGAAATAATC ATTTTACAT ATCCAGGAAT TTGAATTACT CTCAAGTCAC
57001 CTATTAATTA CAAGTCATTT TGAATCATT CATTTTCTTT GTGTTTGCTT
57051 TATAATGTCA TTTTAGATTT CATGCATCAT AATCAGCCAT CAAATAATTT
57101 AGTTAATACT TGATTTTCC TCAGTTGTAA GAAGTGCTGT GTTTAAATTT
57151 CATTCAGAAT GTTTCATTT ATCTGAATTA ATATCTGTTA ATGTATGTAA
57201 TATACACATA TTTTAAACAT GCATGTACTT AAATTGATTA TAGGGACTTG
57251 GTAAAATTAC TTATTTATAG GATATTTTAA ATATAATCAA GGATTTTTTA
57301 AATCTACAGT TCCCATTTGA AAGTAAAAGT AAGTCTTTGT TTACTAGTTT
57351 GTTCACAGTA CAAGTAAACT TTCTACCTTT TGGTTAAATG TGAGTGCAGC
57401 CCCCACAGTG AGAAATTGTT ATATTAGAAC TCTAATAGCT ATAATTTATA
57451 GGGATGAATT TCAATGAGTT TGGTTCTAAG AAATAATCTG TTGGTTTTAA
57501 CAACATTTTT AAGTATCAGA TATTCATCTT TACTCAGTAT GTGACATGTA
57551 CTCTCATAGC TTACGTGCTT TTCCTTTATT TGGGGTGTTT TTTATATATT
57601 AATTGGTATA TCGCATATTT AACTTGGCA TAATTACATT TATATGGACT
57651 CTAAACAATA ACTTGTATTT TAATTTTAA ATTTGAAATG CATCTATGTC
57701 TCTGTTAAAA TGCATTTCTT TCCCTTTGCC CAAATGGGGT ATGGTAAGTC
57751 AAGAGAGTCT CTAGTTAGCT CACCTCTCAT TTGACTGGCA GAGTAAAGCC
57801 CTTGTTTCACT AGAATGTGTG TTAAGCCTTC CCTCCCTTTT GTAAAGTTGT
57851 TCTGAACAGA GCTGCATAAA ACCACAGGTA AAGTGTTAAG CTGATTCTAC
57901 TAGCATGTCC TTAGAAAGGA GAGCGGTAT ATTGGCAGGT CCTATTGCCT
57951 GGCGTTTCTG ATCAATAACT CACCAACAAA CAGAAAACAG AAGCCGCACA
58001 AGGAAAGGCA GAACTAAATA AATGGTAATA GCAAACAATA AGCCAGATAG
58051 CCTCTGGCCT CTCGCCACA CCTTAAGGCA GCTGGGTCAG GTGGGATGCT
58101 TTTGTTTGTC TTTTAACGTA TTTTCTTTAC AAATCTCAGC CATTACATAA
58151 TTTGGAAATG GACACAAGGC TAGTTATTAC TAACATTTTT AAAGACATTA
58201 CTGAATGAAT GTGTAAGAAA ACAAAGGTC CTTTTGCCT TTCAGCAGAT
58251 AAGTCTTTTA ACCAAAAATC TCTTGGGTAT TTTGAGATTG TGTCTACTT
58301 CTTTGCTTAT TTAATATTTT CATAAAATTT GCTAGTTACT CTTGCTTTT
58351 TGCATCTCTT CTAAGAGAAA ACAATTGGTG CATATTATTA ATGAGAAACA
58401 CTTCAAGTGT TGGACAATTT TTTGTAGTGG AAAAGAAATG TGAAACTTTA
58451 TGTGTCAGAA TCATTCTTGG TTCAACTAAC TACTAATTTT AAAACATAAA
58501 GTCTTAAATA TATATAAAGT TTATATGGGT AAATATATAT TACATATAAT

Fig. 2 (cont'd 32)

38/124

58551 ATATGTTTTA TATTTATACA TAATATACTA TATATTTATA CATGATATAC
58601 TAAATATTTT CCCATATAAA TAATAAAATG CTCTAGGCAT ATATGTGTGT
58651 GTGTGTATAT ATGTATATAT ATATATACCT TCATAACATA CATATATAAA
58701 ATACTATATT ATATATACTC TAGGTATACA TATATGCCTA TATATGCACC
58751 TATATATTTA TATATTACTA TATAATATAT AGTATATATT ACTATATATA
58801 CTACTATATA TTACTATATA ATATATAGTA TATATATAGT ATATATTATA
58851 TAGTAATATA TTACTATATA ATATATAAAT ATATGTGTGT ATATATATAT
58901 ATGCCTAGAG TGTTTTTAAT TTGTCAGTGG GCTGTCTCTG TAATCTATAT
58951 GAAGAAATAA AATGTAGACG TTATGTATAA TGATATTTCA TCTTGTGTGT
59001 TGGCATCATA GTAATTCTCT TTACATATCT ATTCAGATTA CTTTTCACC
59051 AGCCTAATAC ATTGTATGAT TCCAAAACCA AAGAGAGTAT GGATTGAAAT
59101 GATATTCCTT TTACTAATAC TCAGTCTTGT CTATTTTATT ACCTTTATAG
59151 ACTTCACCTA ACACAAGTCA GGGGATATTT ATCATCATAT TAATACAATT
59201 TTACTCTGAC CTAAAAATTA TGCAACTGCT AAAGGAAAAA TCAGAAACCA
59251 ATAACTGTC ATTAACAACC CCCCTGAAAA TCCATATTTT TTAAGTCA
59301 TTTTATCAAG TCTCTCAGAC AAGATGTGAT ACCCTATAAG TTTAATCAGT
59351 TTTACTTTCC ATTTTCTCTT CATTAAGGTG ATAAAGATTA TCATTAGTAG
59401 AAAAATTTTC CCTTATTTGC CTCCTTTTCC ATTTACCCTA TTGAGTGAGA
59451 AATTTAGCCT CTCATAACTT CTAAAGTAGC AATGTTAATC TGATAAACTA
59501 AACCAAGGTG AGATAAATTT AAGACAATAT TTTTTTCTT CAACTTTTAA
59551 GTTCTGGCGT ACATGGGCAG GATATGCAGG TTTGTTACAT GGGTCAACAT
59601 ATGCCATAGT GATTTGCTGC ACAGATCAAC TCATCGCCTA GATATTAAGC
59651 CCACCATCCA TTAGCTATTC TTCCTGATTC TCTCCCTCCC CTAACCTCCA
59701 CTGACAGGCC CTAGTGTGTG TTGTTCCCCA CCATGTGCCC ACGTGTCTC
59751 ATCGTTCTAC TCCCACTTAT AAGTGAGAAG AAGTGGTGT TGGTTTTCTC
59801 TTCCTGTGTT AGTTTGCTGA GGATAATGGC TTCCAGCTCC ATCCATGTCC
59851 CTCAAAGGA CATGACCTCA TTCCTTTTTA TAGCTGCATA GTATTCCATG
59901 GTGTATATGT ACCACATTTT CTTTATCCAG TTTATCATTG GCATTGGGT
59951 TGATTTCATG TCTTTGCTAT TGTGACTAGT GCTGCAGTGA ACATAATGCA
60001 TGCAGGTATC TTTATAATAG AATTATTTAT ATTCCTTTGG GTATATACCC
60051 AGTAATGGGA TTAGTGGTTC AATTCTGCT TCCAGATCTT TGAGGAATCA
60101 TCACACTGTC TTCCACATTG GTTGAACATA TTTACTCTCC CACCAACAGT
60151 GTAAAAGCAT TCCTTTTTCT CTGAAACCTC TGCAGCACCT GTTATTTCTT
60201 GACTTTAATA ATCACCATTG TGACTGCTGT GAGATGGTAT CTCATTGTGG
60251 TTTTGATGTT ACCCTTTTTT TTATATGTTT GTTGCTGCA TGACTGTCTT

Fig. 2 (cont'd 33)

39/124

60301 CTTGTAAGTG TCTATTCATA TCCTGTCTAT TCATGTCTTT GCCCACTTTT
60351 TAATGGGGAA GTTTGTTTTT TACTTGCACA TTTGTTGAAG TTCCTTGTAG
60401 ACTCTAGATA TTAGACCTTT GTCAAATGGA TAGATTCCAC AAATGTTCTC
60451 CCATTCTGCA GATTGTCTGT TCACTCTGAT GATAGTTTCT TTTGCTATGC
60501 TGAAGGTCTT TAATTAGATC CTATTTGTCA ACTTTTGCTT TTGTTGCAAT
60551 TGCTTTTGA GTTTTTGTCA TAAAATCTTT GCCCTTACCT ATGTCTTGAA
60601 TAATATTGCC CAGATTTTGT TCTAGGGTTT TTATAGTTTT TGGATTTTAC
60651 TTGTAAGTCT TTAATCCATC TTGGGTAAAT TTTTGTATAA GGTATAAGGA
60701 AGTGGTCCAG TTTTAATTTT CTGTATATGG CTAGTCAGTT CTACCAGCAC
60751 CATTTATTA TTTGTTTTTC AGTTTCCCCA TTGCTTGTTC TTGTCAGGTT
60801 TGTCGAAGAT CAGATGGTTG TAGGTGTTTT TCACTAACAT AATCATAACA
60851 TACATTTTCA TGAAAACAAC ACGACTCAAA ATGTTCTTTA GTAACCAGTT
60901 ATAAGTTTTT TTGTGCATAA TTACAACTG CCATTCTAAT CATAAACATT
60951 TTGTGGTTAC TTATAGCTAG AAAATGTGAG TAATATAGTT TATACAGCAT
61001 ACTCTTTACA ATCCCGATT CTTTGTCAAA CTTTAATTCA TATTAAATTG
61051 ATAAAGTATA CACAAAGGGT AAAGGAGAGT AATTTTCTTC AAGTTTCACA
61101 TTTAAGGATT CATAGTAGAA TGATTAAACC TTACATTTCT CCACTATAAG
61151 GAGAATTAAA ATGGAAATAT TGAGTAAAT CTTACATTTT ATTTAGTAAG
61201 TGCTAATAAA GGGTTTCTGC CATAATTTTC CTTATTTTAA AAGAAAACAC
61251 ACAATTTTAG TTTTAGGTTT TAGTAACCAA TTTTATGGGC ATAGTGGGAA
61301 TATTCTAAC AGGTAAACT GAAGTGACCA TCATGGGCAT ATATATATAT
61351 TTTAAATTCA CATATATGAA TACTATACAG TAAAACTAA CTTATGCTAC
61401 ATACCACATG GATGAATCTC AAAACCCATG TAAAGCAAAA GAAAACCACA
61451 AAAGAATCAT GCCATTTGAT TACACTTGGG TGGTTTTTAA AACAGGCATA
61501 TCTAAACATA GTGCTTTAAA GTGTAAGCTT GGGTAGGAAA AACTATAAAG
61551 AAAAGCAAGA AAATAATTAC CACAGAAGTT ATGTAGAGGT TATCTTTGGG
61601 GAAGGAAGAG GGAATAATAA GAGAGGGACA AAGAAGAGCT TCTTGGTTCT
61651 TGAAATGTCC TATTTCTTGA CTTGGCTGGT GAATGCATGA ATGTTCACTA
61701 TGTGATAAGT CAGGGGGCTG TTTTCATTTT GTTCACTTTT ATATATGTGT
61751 GGATTTTTTC ACAGTTGAAA AGGTAAAGTT CAGGTGTGGT GGCTCACACC
61801 TATAATCCCA GCCAACACTT TGCGGGGCCA AGGTGGGAAG AATTACTTGA
61851 GGCTAGGAGT TGGAGAGTAA CCCAGGCAAC AGGTGAGGC ACTGTCTCTA
61901 CAGAAAATGA AAAAAAAAAA AAAAAAGTAG CTGGGCATGT TGGTACATGC
61951 CTATAGTTCT TGCTACTTGG GAGGCTGAGG CAAGAGGATC ACTTTAGCCC
62001 AGGAGTTTAA GCCTGCAGTG AACTAGGGTT GTGGCACTGC ACTCCAGCCT
62051 GGGTGGCAGC AAGACACTGA GTAAAAGAAT AAAATAAATA ATTAAAAGTT

Fig. 2 (cont'd 34)

40/124

62101 AAAATATAGG AAAAAATGAG CATAGCCTTA TGCTAATTTT TCAGTTACTA
 62151 GGTCTGATAT CATCACATTC CTTGCTTGTC ATTGAAAATT TTTTAAACTA
 62201 TGATACTTTT TTTTAGTGGT ATTTATCCAA TTAAATCTGC TAACAAATTT
 62251 GGTGTATAAA TCTCAAGGGT AAGGGTATGT GGAGAGTGGG TGTGTTTGTG
 62301 TGAGAGAGAG AGAGAGAAGA GGGGGAGGAG AAAAAGAAGG AAGAGGGAAG
 62351 GAATGGAAAA AGATAATAAA GAGTTGTTCT GATAGATTAA TCTTTAGTAG
 62401 ATGTATTCCC TACAAATTGT TTTTCTCCAT ATTGCAGTGT CAGGTAAAGA
 62451 AAGGCATCCC AGGATGAATT CAGAGCTAGG AACATGCACC TTTGTATCAT
 62501 AATGCTAATG GAAGGAACAT GTACATTCTA ACTGTTACCA ATAATGGAAT
 62551 ATATTTCCGT TATTAAGTAA TAAGCTTTAA TTCTTTGTAT TTTTGTGATC
 62601 CATTTGATAG TAGGTGCCTC AGCATTTCCA CTCTGCTATA AGTACATGGA
 62651 GATATATTTT ATTTAAGTCA TCTTATTCAT GTCTTTCAAA AAGAAATTCA
 62701 TTTTGGCCCA AGGATTTCCA AATTTTGCCC CATATATAGG TATAGTTTAT
 62751 TATAGACTTC GTTTGCAAAA TATTAAATCC TTATATCCTT TTAGGGACAC
 62801 AATAAAATTT TATAAGTTTG AGATAATGTA CTTCAGTTC TACCTCAGGC
 62851 CGTGGTGAGA GATTGAAGTG CCTCTTCATT TTAACATTTT GGGTTCAAGT
 62901 TGTTGCATAA GGGCATGCAA ATGGAACTG GCCTATTTTT GAGCTTTAAT
 62951 AAAATCGTCA AATACTTCTT AATCTTAAGA GTTATAGTTA TGTACTACAA
 63001 TATGTATAAT TCTCTAATAT TTAACAACAA ACCTGAAAGC CACAAAAGCT
 63051 TACTGTGAAA TAAAATGTGA TGGAATATTA TTTCTAACTG GCTTACCTGT
 63101 ATTTCCCTTCA TTGAAGGGAA TATGAAGTAG AAAAGCCCTT TTATTGAAAA
 63151 GAGTTTGGAA AGTAAAGATA ACTCTTTTCA ATTCAATTCT TTGTAAGTAG
 63201 AAAAAGAGTA AAGATAATGT TTAGCTGTCA GCAGATGTCT GACACTTGAT
 63251 GGAGCGTATC ATTACAATAG AGCAGCTAAC AATATCTGCA AAGGTCATCA
 63301 TGAAAGTATA AAAATGAGGA ATATTTGTCC ATTGACCATT TCAGTGACCT
 63351 CTTTTTGGGC TTAAAGTCTA AAAATCTTGG CAGATCAGAA CTTTATATTC
 63401 GGCATTTTGA GTGTCAAATC TCTACATGAT GTGCAAGTCA GAAGGAGTTA
 63451 TTACTTGCAA AATACCATCT TCTTTCAGAA GTTAAACTCA CATTAATGC
 63501 CAGGAGACTG AAACACTGAT TTTAAGAAGA CAAAGTTTAG AAAAGATGAA
 63551 TGAAAATGTG TGTTAAAGAA GAGTCACCAG TCAGAGCTAA CTATGATAGT
 63601 CATAGTATTT AAAGAGTTGG AACACATGAA ATTAAGCATT TTGTAAAATG
 63651 AAGGCTTTTC ATCCATCCAC ATAAGATTCT GACATTTAAA CTATGTTTCT
 63701 TCCATTCTGT TCACAGGCTC ACCTTGTAGC AGCTTTTGAA AAGAGCTTAG
 63751 GGAATATGAC TGGCCGATTG CAAAGTCTAA CTATGACAGC GGAACAAAAG
 63801 GTATGTTTCA AAATTGCCAC TGGAGACTGA AAGAAGACAG CAAATTGCAT

Fig. 2 (cont'd 35)

41/ 124

63851 AGGATTCTTA AATAATACCT GAAGCTCCTT AAAAATAATA TTCCAGGCTG
63901 AGTGCAGAGG CTCATGCCTG TAATCTCACC ACTTTGGGAG ACCAAGGTGG
63951 GTGGATCACT TAAGGTCAGG AGTTCGAGAC CAGCCTGGCC AACGTGGTAA
64001 AATCCCATCT CTACTAAAAA CACACACAAA AAATTAGCTG GGCATGGTGG
64051 CGGGTACCTG TAATCCCAGC TACGCAGGAG GCTGAGGCAG GAGAATCACT
64101 TGAACCCAGG AGGCAGAGGA CGCAGTGAGC CAAGATCACA CCACTGCACT
64151 CCAGCCTGGG AGACAGAACA AAAAAAGAG TAATAATAAT AAAATAATAT
64201 TCAATTCTAT ACTAAATTAA AACAATGATA ATACCTTTCT TTTTCAATTT
64251 TAATTTAAAG ATTTTATCAG TTTACTCCAT ATTGGAACAC ACAAAGGCAA
64301 ACAAATCCT TGCTGGGCAG TCTATTAATT TACTTCTGGA TGGAACTAGT
64351 AAAAGAATAC TGAATGTTAA GAAAGAGAAA CAGTCACATA AGAGAATATT
64401 CTGGGGGCAA ACTGTTATGC AGTTGACAAG AATCACACTT TGATAAGAAC
64451 TTTACAAAT ACATGGTCAC TAAATCCAGC TATAGGGCAT GGC'TGTAGGC
64501 TAAGACACAC AGGAAGGATG CCTGGGACTC TGCCAAGTAA GGGACTTCAG
64551 GTTACAGCAG CTATGAAACA AAGGCCAATC CTGTGTAATT TTGAAATAAC
64601 AAGAACTAGT TGCCATCTAG GGATATCACC TTTGAAGAAA AGTCATTTGT
64651 TATATCAAAA TACTTAAAT GAACCTAAAG GATTTTATGG TATGAAAGAA
64701 GGTATACCAA AAAGAAAGGA ACGGAGAATT TAGTTCACGA AGACAAATGT
64751 ATTA AAAAGG TCCATACTGC ATAGAAAGCC TGGTCACCTT TCCTGTGATG
64801 ACCAGTTAGC TTACTTCTCT GCTGTTAGTC CAGTGGCCTT AACTTCCTTG
64851 GATAGGTATC AGAGATAGGT GAAACCTATA GAATTCTATG GAGTGTGTGT
64901 GTGTGTGTGT GTGCGTGCCT GTGTGTGTGT GTGTGTGTAT GAAACTGTAT
64951 AATGTGCATA AATGATCAGG TGTCCAGAGC TTTCATCTAA TTCTCAAAGA
65001 GACCCATTAT ATCAGAAGTT TTGGGTATTT TCAAGAATGC GTTCCTCTAT
65051 CTATCCATAG GAATGGCTTC AGTTTTGTCT TTAGATTCTG TAAGTTATGT
65101 GATTAGCTTT ACAAAGTAG TATGTATTAC CAAATTTTGT CACTTTACAA
65151 AAGTTTATTT TTA AACAGA ATGAATAGTT CAATGAAATC AAAAGAGTAA
65201 ATCGAATATT CTTATAATTG CCAAGTATTA TTAGCACATT GTATTCTCTC
65251 TCATATTCTC CGTATACCCT GCCCGTGAGA GAGAATATTA TCCATTCCCTG
65301 GAAAATCTGT TCTAGCACAG CTAACAACT CCTTTTGAAA CATAAATTTT
65351 CCTTCTTTC CTCCCTCCCT CCCTCCTTCC CTCCCTTCCT TCCTTTTTC
65401 TTTTCTTTC TTCTTCTCTG CCTCTTTTCT ATCCTTCTT TCTCCTCCCT
65451 TACACCCTTT CTTCCTTCTT TTCCCCCTCT GTCTCCCTCT CTTTCTTTT
65501 TGCTGCAGCT TGTCACCTCA CTATGTAATA TAAGAACCCA GCAAATAGAA
65551 TTAGAAGGCT TTTTAGAGCA GCTGACGGGA AAGAATAAAA ACACTGGCCC
65601 CCAGTATTCT TGAATGAGAA TTCTGGCTAT GTCTGTTAAA AGCTGGGTAA

Fig. 2 (cont'd 36)

42/124

65651 TCTTGAGCAA GTTTATCTAA CCTTCTTGA ACCTCAAATT CACCTTCTTA
 65701 AAAGTGGGGA TGATAATGAC TACCTTGTAG GATCACCATG AGGAGTAAAT
 65751 CAGATACTGT TATCATGTCA CATGCTAGGG GCTACCAAAA AATATTACCT
 65801 TCCTTTACAT TTCTCTTTTT CCCTTGAAAA TTATAAGATA ACACCAAATT
 65851 CCTCACTGGG CATATACCAA GCATATTGTT GGAAATGAGT GTTAGAATTT
 65901 AAGTCTCAAT ATCTTTAATA AGTCAAAATT AATAGAATTT TTGTCCTCCA
 65951 CCCAATATTT TCTTGAAGTC TGTATATCT GTAAGTGAAT TTTCTCATAG
 66001 AACATACAG AGAATTTTCT CATATACATA TAGAAAAAAA TGTAGAGGTA
 66051 TGTTAATGTA TAATGCCTAT GATTAATGCC TGAATATTTA AAAATAATTT
 66101 CTATAACATA AGAGATTTTA TAATGTGTCT ACATAATCCT TAAAATAACA
 66151 TTGCCAAAAAT TATAAAATTT TCTCAGAAGA TATCAGAATG TCTCATATTG
 66201 TCCTTATCAC TTTTTTAACT GAAAATAAAA TCACTTCTTT TTGAATTGCA
 66251 AACTGTATAC ACACAACAAT CATGGTTAAC TAGTTTATTA ATTTGAGATT
 66301 ATAACCTGCC TATTCTCAAA GTGATATTTA AAAGCCTATA AAATTATTTG
 66351 CAATGTGAAA TGGTATAATT CAAAGACAGA ATCTAATTAA AACCAGTAGA
 66401 ATAATGTATA TAACAATATA CCTCAGCCTA GATAATTACT ACTGCAAGGC
 66451 ACTGAAATGA ATTGAATTTT AAGGAAGCTA TGGTACAAAG GGAGATTGTT
 66501 AGGTGTGTTT TATTCTCAT TTTCTGACCAG GAGAGCATAA TTTAGACTGA
 66551 GGAGAAAACCT CTTTGGCACT AAATTCAAGG ACGAATTTAT TGCCAAGGTT
 66601 TTAAATTTGG GGTATGGAA TAACAAAAGA CAAAATCACT GTTCAAATAG
 66651 ACATTTCTCT AAAAGCTAAG GGCATAACAT TTAATCATAT TTCACTAAAG
 66701 GCATTTCTTC AGGGAGCTGA GATAAAAGGG TATATTGCTC TCTGGTGATT
 66751 CAACAATCCT GAGAAAAGGC TTGTGAAGTA TAGAGCAGAG ATTCTTAAAC
 66801 TCCCTTCCCC AAGTTATAAG TTTCAATTTGT CTATATAGTC ATTCATCAAG
 66851 TTTATATTGA ATTTGTGCTC TTCTAATGAC AAAACAGTAC AGACAATATA
 66901 GATATAGAAT GATAGATATA GGTCTATATC TATAGACATA CCTATCTACT
 66951 AGAACTCTAA AAGCATATTA TACATGTATG TAATATTCCT CATGGAGTTT
 67001 ATATTTCTCA TATATATCTC ATATATATGT ATCTCTTTAT CATGGAGTTT
 67051 ATATTTTAGG AGGTCACAGA TGATAATAAA AATATAATTA AAACAGGCCA
 67101 GGTGTGGTGA CTCACACGTG TAATCCTAGC ACTTTGAAAG GCCAAGGCAG
 67151 GTGGACTCCC TGAGATCAGG AGTTCAAGAC CAGCCTGGCC AACATAGTGA
 67201 AACCCCATCT CTA TAGAAA CAAAATTAG CCAGGCCTGG TGGTGGGCAC
 67251 CTGTAGTCCC AGCTATTCAG GAGGTGAGG CAGGAGAATC ACTTGAACCT
 67301 GGGAGGTGGA GGTGTCAGTA AGCCGAGGTC ATGCCACTGC ACTCCAGCCT
 67351 GGGCAACAGA GCAAGACTCT GTCTCAAAAA AAAATATATA TATATAATAT

Fig. 2 (cont'd 37)

43/ 124

67401 ATATAATATA TATATAAATA TATATATTAT ATAATATATA TATAAATTAC
67451 ATATTTATAA ATATGTAATT TATATATATA ATATATAATT AAAACATATA
67501 GGATTTTCAGG TGATGATAAG CACTACTGAA AAAAGTAAAG CTGAGAATGA
67551 GGATACTGAG AAGCTGGTTT GGAAGCTAAA ACACAAAGTA ACAAAGGCCA
67601 AGGTGGTTAC ATGTTCTTGA TTACATACTT TAAAAATGGA TAAACTAAAT
67651 TAAGACTCAG ATTCTAGTCT TTGGGCTTCA CAGTGTGATT TTCAGCAATC
67701 ACATGGCATT AATAGCCTGA AACTACATCA AAATTGTCAT TTGATTTATA
67751 GACCAAAATA ACTCCCTTGA ATAGAGAGGG ATTCACCTCT AACACTTTTC
67801 CTATTTCCAG ATGCCAAATA ACACGGAATC TCTTGCCAAA TTTGTGTGGC
67851 AGAAGACTGG TTTTATATAC TTATAGCCTG GTAAGAAAGA AAAGACATGT
67901 ATGAATAACT TAGAAGGCAG AAAATTATCA TGCTATTAGA CTCAGTACAA
67951 TGTCATGTGC ATTCTCAAAG GAAACATCTG CAGAGGCAGG AGAATTGCTT
68001 GAACCTTGGG GGTGAAGGTT GCACTGAGCT GAGATCATGC CACTGCACTC
68051 CAGCCTGGGT GACAGAGAGA GACTGCATCT CAAAAAATA AAAATTACAA
68101 AAATAAAAAA TAAAAAATAG TGATCAATCT GGCAGCATTT TCTGAAAGTT
68151 AAGCAGTATT CCCAATAGCT GCTAAAAGAA GACATGTTAT ATAATACTAA
68201 GTCTGTAAGT AGGTAAAAAT TAAGAGAATT GTTAATGTGC TTGCTGGGGA
68251 GTGAAATTAT CTCTAGGCAT TACCCTATAC CTAACCTAGG ACTCAGTAGA
68301 CTATGATATT GCGTAGTTT GACCAAGAAT TTTATCCTGA TTTCAGATCG
68351 TTTTCTCTTC ACCAGCACTT CTTCAACAGG ATTATATGAA AAAAATTAAA
68401 CCTGATGCCC TGAGGCATCC ATTATATGTG CTGAAATAAC TTCTTTTCTC
68451 ACCATCTAGA ATGGTACTAG CTATGTACCA CTCTTGTCAG AATCAAGGAA
68501 ATTGCTACTC AAATCATTGT GCAGCTTAAT TTTCTCACAG AAGGCCAGTT
68551 GAGAAAGGCT CAACTTCTAG GAATCCAGCA AACTATATTT TTTATAAGTA
68601 ACATTTTAC AGAACTACTT CTAAATCCTT GTGTTCAAAT TACTAAAGC
68651 TATATTCACA GCTAAATATT TCAGAATTTA AAATTTAAAA GACTTTCAAA
68701 TTAGTTCCT GTAGCTGTCA TGCCAAGGCA ATTAGAACAT ATGTTAAGGT
68751 ATGAGGGGTT TTTCTTGTTA GAAGGTCAGA GCAGGGCAGA GAAGTAGCCC
68801 CTTGTATGAG TGATGAAGCT CAGATATTGA CTCCTATGCT AACCATAAAG
68851 CCTAGTAGTT TGCTCATTTG TTACCTCTCT GAAACATTTT TTTGGGTGAC
68901 TACAAAACAG GAATTGAAAC CTTCAAAATA AGGGAATTG AAACCAAATC
68951 TTTGAAAATA GATAATGCTG CAACTAAAAA TTTAGTTGAA TAAGATTTTT
69001 ACATTAATCT TCCCTAATTT ACGTTATGAT ATTTGCCATC TAGAAGTGTT
69051 TTTAAAAAAT ATATTGCTGG AGTCAGATGA TGCATCCATT AATCTTTGGG
69101 GCATAGAATA ATGTGAATCT AAAATTTTCA AATTATTTAC ACTACTGGTA
69151 TTTGGTCAAT GTAATTTATT TGAAACTAGA TGCAATAGGG ATGGCCAGGT

Fig. 2 (cont'd 38)

44/124

69201 TATTTTCAGTA GAACAACTAG CAAGACTTCA GATGCATGGT GGAGTGGGGA
69251 AAGGAGGACC TGTTTAAGGA AACTAGAGCT GGGAAAGTGTG AGATTAACTT
69301 AGTGCCAATG TGAGGACCTA AAAAGCAGAT GTGGTGGAAG ATTTAAACAG
69351 GCTTGCCTAG AAGGTCAAGT TAGTTGATGA CACTTGATGA GATTGTCCCA
69401 AGCTTTGGGA TTCTCAACAA AGTCTTTGTT AGTGAGAAAT TTGGAAAGAG
69451 ATCAGGTATA GTTAAGAAAC TGGGTTGGAA AGGCCACCAG GAAAGGCGAA
69501 TATTCTGACA CAAAATTTGA TCATTTTATT TGAAGCATT TCAAGCCTGA
69551 CCTGAACGAA TTGTTTAGCC TCAGATACAT GCATAAACT GTGAAAAGAG
69601 ACATTGACTC AATTTAGCTT CTTTAACATG AGAACTTTC GTGAAAACCT
69651 AGAACTTTAC AAGCTCAGCT GGTGTTGGGG GCATCATTAT CTTGAATAGC
69701 TCCTGGAGG AAAATGAAAT CTTAGTTTGG TTCTCAGGTT TTAATAATATC
69751 TATCATTTTT GAAAAGTGTG AAGTAACAAA ATATGATCTG ATTATCTTAT
69801 TCCTAAATC CTTTGAGAA TTATCCCAGC CTCAATCTTC TCTTTAGTAT
69851 TTAATGAGAA TAAGAACTG GAAATGACTG AATTGGAAGA GTAGACTTTA
69901 AATCCATATC TTGATGGCAT ATACATTTTT CAGTTTTTTT TCTAAATGAT
69951 TAATGAGGAT TCTCAAACT TGAGTATCTT CTATGTTTCC CTTCAACATA
70001 AAGAAATTGT ATGAAAATAT TTTAAAAATT TCTAATGATT TTATAGTTAG
70051 CTATCTGGG AATTCAATTC TAATCATGTA CCTCATCCAA ACTCCCCACT
70101 ATGGACAAAA ATAAATAAAA AATTATTAGT TGCATCTGAA GGCCACATTA
70151 CAATTTCTAT GCATTATAGA AACCTGAGAA AATGTATCTT AAAAAATAAA
70201 TGTGAACAAC TAACCATAAT TATGAAGAAG AAAAATGAAA ACTAGAAATA
70251 AACTATTGAA AAATGTCTAT GTATCAGTTA AGTTTTTATT TTAATAATTCT
70301 TTATGTTTAT CTCTATAATA CTATTGGGAA AGAGAGAAAAG GAAAACCTGA
70351 CTTTGTCTC ATCCAAAGGA GGTGATTCCA CTGATTTAGC CAAAATAAGA
70401 CTTCTGGTT ATAATAAATA ATAAAGTTTT TGATGTTTTT TATATGGTAC
70451 CCCACTCACT AGGTGATCAG ACACCCTCCT GCAAAAAAAA AAAAAATACG
70501 TATGCAATAA AGTTAAAGTT TTATGTTATT CTTTCAAGGG GAGAAACATC
70551 TGTTTAACAC AGACCAGAAT ATTTCAACAA AGTCATCCCA ATATTTATGG
70601 AGATCATAAA TCAAGCGAAA AAATATATTC ATCAACAACT AAACAACTA
70651 CATTAAATAG TCTCAAAGCA CATTTTCACT TTTTTTCTGA CAGGAAAACA
70701 GGTTTCACAA GTGTGGAGAC ATTTTACCAT GGCTTTTAAC AGTGAGGAAG
70751 GATGTTTAAA TAAAGGGAAA AATTATATGG AAAGCTCAGA GAAAAGAGAT
70801 GGGTGTGGCT TGAGTGACAA GGTGAGAGCA GATCTCATT ACTGAAATGA
70851 GAGAGAAGGA AGGAATTTTG CAAATATGGA AAGATAACTA GTGCAAGTTT
70901 GAACAGATTA TGTCAATCAA TGTAGAATTT GGCTATCTTT TTAATCAAAG

Fig. 2 (cont'd 39)

45/124

70951 AAGACTATGG AATATTTTAT AGGTGTTTGC TTATACTCAA AGTTTTAAAG
71001 AAATAACAGT ATGAATTTGG TTGAACTAAT TTTTTCATA GATAGGATTC
71051 TCCCAAGTTA TATAGCATAT ATATTCTTA ACTAGTTATT CTTCCCTTTA
71101 CATATATTGT GCCACATTGA GTAACAACTA ACCTGCTAAT AGCTATTGGT
71151 TTTTAAAGA TAATTAATAT TAGAAAAGTGA TCATTTTCTT GTTTCATATT
71201 AAACATGATA TTCTGAAAAA GCAACATTGC CTGAATGTTT TACATTTTAT
71251 CTTTTTGAAA ACAGGTTTTA TAAGAGATTT CTTGTGAAAA GCTGAACGTT
71301 CTGACACTGA AATAAGTCAG CTAACCTAAA GCTAAGCTTA ATTTTTTGAC
71351 ACTGTTGGCA TGAGGTCTCA TTCCCAATTT TTTCATTTAA AGCCACAGGC
71401 AAATGTTTTA ACAGATTTTA ATCCGTAAGT CAAGCATTAT TGATCTTAA
71451 TTTAAGGATA AAAACCTGAT TTTAATTAGA ATTTAATATG CATTCTAGTA
71501 TTTACGTTGT ATAATTAATA TTTACATTCC ATGATTCCAC TATGTACCAT
71551 TTATTTCTTT TTGAATAAAT TTCCAGTAGG AGCAGAATAA ATTTTCAGTG
71601 AATATTTTAT TTCTTGGGGG ATATTTTTTAA ATGGAATAA TATTAAGTTT
71651 CGGTAAAATC TGTGCTAAT TTGGCAGTGG ACAGAATATA AAAATTGGAG
71701 AGACTGAGTC ATTATGATGA ATTGGGTCTG ACTTTTGTCA TGACACTGGA
71751 AATTTCCAC AAATATTATA TTCTTCTTTT ATAATAATA TAGTCGAAAT
71801 GAATTGCAGT CAAGTATTTG AAGACCCATC TATAAATTTA GCGGTTACT
71851 GTTGATTTTTT CATTATGAGA GATTCTTCCA CTCATAAGCT ACTAAAAGTA
71901 CATAAAGAAG GTCTGGTTGT TTGTTTTAAA TGTGACTGTT CTCTATCAGG
71951 AAAATGTCAG GTATCCGATG AAAATAGATA TATGAGGTGC CAGGTATCTA
72001 TTCCAAACTT GGATATCACT TCAATTAGCA TCATCTTTT TTTTTTTTAA
72051 AGTGTCTAAG GTTAGAATAG TCACCAGATA TTCCCATGTA TGAAGCAATT
72101 TTCTGCAAAG GCCGCTGTGG ATGATCTTTT TAAAATATAT ATTCTGGGAG
72151 ACATTGAGTA AAGAGAAATT ATTTACCAGA GAATGAAGAA CCGAGGCCCG
72201 ATTCTTTGGC TTTCTGCCAA AGATGCTGAA GGCAGTGATG AATGACAAAT
72251 ACATTACCAA GGAATTCTCC CTCTAAGAGG CTGACAAAGA TCTGATTTTT
72301 AGGATTATAT TACCACCAAG AAGATACCCC TTGTCACTGA GCTTCTAATG
72351 GAAATATGGT CTATACTGAA ACAATTCTCA GTTCTTTTTC TTTCTATCTT
72401 TTTTGTAGTT ATTTTATCTT CCAAAAATGA GTTATTTCTG ATAAAATAAT
72451 TCACTTAAAT AATTATGAAA GTTCAAATTT GTGCAAATAT TTTTATTGGG
72501 ACATCTTAAA ATTACTCTAA ATTCAAAAAG AAAATATATG CTTTATTAAA
72551 ATTTGATCTG TAAGCTGCTT TGTTTGTAAT TTAATAATTA TATAAAAATT
72601 GTATAATACA TATATTTTAT TTACTTTATT CCTGTGTTGC TTTGGCTTGG
72651 TGAGACTAGG TCTCCACATT AGGAGTTTTA CTGAATGAAA AAGTATCAGA
72701 ATGTAACATG ACTTTGATAT GGCATCAGAA TTTAATAAGA TGACATTTAA

Fig. 2 (cont'd 40)

46/ 124

72751 TAGGAATTAG GGGTAAGTTC CAGGTTTTAC ACTTAAATAC AAATAATCAA
72801 TTTTGCAGGC ACAAATACT TCAAACAAAA TCTGAAATCA TTCATTTGAC
72851 AAAACTTCAG GTTTGCAGTT GACAATAAAT ACAATACAAT GCAACAGTGC
72901 AATAGTGATA TCTAAATATC TAATGTAATC ATAGGTAATA TTAGTAAGTG
72951 TGTATCTGA AATGAGTGGT GTGATATCCT GCTTTACTTT GTACTGGTGA
73001 GTTCTGGGTG CCACCTTTGA AAGGAATAAA GACTATTCAT ATCTCTTTTA
73051 TAAGACAATA AGAAAAACAA ACAACAAAC AAACAAAAAA CCACCTCCTT
73101 TACTTTAGCT GAGAAAGAAG TTATTAGGTA CAGCTTGACA AGTTCAGCTA
73151 AGCATCCAAA TCTTCCAGGA GGTGTTACT ACATAAAATC AAACCTTTTT
73201 AATTCAACTA TGAGCAGGGA GATTTTATTT TTCTTCGGG TACTAAAGCT
73251 TCCAAACTCT GTTTATTCCA CAGGAATCTG AACTTATAGA ACTAAGAGAA
73301 ACCATTGAAA TGCTGAAGGC TCAGAATTCT GCTGCCAGG CGGCTATTCA
73351 GGGAGCACTG AATGGTCCAG ACCATCCTCC CAAAGGTATA TTTAGAAATC
73401 ATTCATTTT CACCCAATAT AATAGGCATC TATTTTATTT ATTAATTACA
73451 GTAGAACTGC ATTTACTCAG TGTCCTGTG CATTATTAAT ACATACTAGT
73501 TGTATTAATA GTTGATTAA TACATACTAG TAGTATTAAT ACATACTACG
73551 TTGGTATTAA TGTGATCAGA ATCCTAGAAT TTAGAACAG TGACTTCCAT
73601 TATCAGATAA TTTTAACT GATCTTAAGA AATTGGTTC TATAGTTGTA
73651 TACACATCTC TCTACTTGAT TCAGTGGAGA TGGAGATGGA GTGGTTGGTT
73701 AATACATGCA TATCTGACTT CAGGCAAAAC AAACCCATTA ATGAGTATGA
73751 TAATCTAGAT CTGTATTTAA AAATGAAATA GTCAATATGA TGATATAGTA
73801 AGCAGTGGGC ATTGGGAACA ACTTTTCCTG GATGGAGGCT ATAAAAAGGT
73851 ACATTTCTG TAGATAATTT TGAAACAATA AAAACAACGG GTGAAAGGTA
73901 GCTCTGTTTT AAATTATTCC TATGCTTAAG CAATTCTAAA CAATGAAAGG
73951 GGTATTTCTG CCACTGCCCC TACCCCTGGG TTCACCACTG AAGAAATGCT
74001 CATTATTAAT ATCGTGCAT TTTTTCCTT TACATTGGTT CTATTTACTC
74051 ATTTCTGAC ACTTTTCAAT GGCCTTCAGT GAGCTCAGCT CTTTCCAGC
74101 TTAAAAATC CTGTCCTAAA ACATGAATGC CTTATTATCT CTCTTTTCAT
74151 TTCCAGAAGA ATTCTGAGAA AAATTTTATG AAGTCTTCA ATGTCTTCAG
74201 CCATCTTTAG ACCACTGGAG TGTAGCTCCT TTTCCCTCCA CTCCACCAAA
74251 ACAATGCTCT CCAGGATCAG CAGAACTTA CATGACACTA AATTCAGTAA
74301 AACGTTTATA ATTCTTATTG TATTAGACAG ACATGGAAAC AGCATTTGAT
74351 GCTGATATTC ATTTCTTCTT ATGTGAAACA TCCGGTTTTT CTAATGTTG
74401 TGACATCATA CATTCTTGGT TTTCTTCTG TTCCTTTGAA ATATTTTTTC
74451 AATATTTCTT TTGTAAATTC ACTCTTTGT ATCCATTGT TAATTGTTGA

Fig. 2 (cont'd 41)

47/124

74501 TATCCTAAGC TCTCTTCCAT TATGATTCTA TGCATCCTAT TTAAAATATA
74551 TAGAAAAATCA TCTCATACTC TAGCTGTAAT TTTTATTAAT GTGCTAATAG
74601 CTAATAACTG TCAAATCTAG GTCTCCAGGC CAGGCTCTGT ATATCCAGCT
74651 ACCAAGAGAG AACTCCACGT GGATATCTTT GGATGTCTGT TTTGCATCTT
74701 AAACCTAACT TCTCCAAATT TGCACTTGTC TTCTGTCTCA GACCTGCTGC
74751 TCCTTCAGTG CTCTTTGCCT CAGTAGATAG CACCACCATC CTTCCATTTA
74801 GCCAGAAATC TAAGTATTCT TCATAACTCC TCCTCTCCTC ATTGAATAAA
74851 TTACCAAGAT CCGTTGATCC CATTCCTTAA ATATCTCTTG GATCTGTTAA
74901 CTTTCTCTCTG ATTTTACTCT TGCCATCCAT CACCTCTCTC CTGAACCATG
74951 ACCACAAACC CCTAAATAGC CTTCTCTCTC TTAATCTTAT CCTGCTTTAC
75001 ACCAGTCTTC ACGCTGAAGC CAGAATAGTC ATTAAGAAAC ACATCTACAG
75051 GTATCCCATT CATTGCCTTT AGAATGGAAT ACAGACTCCT CAGCATGACA
75101 TAATCTCTCT TCACCAGCTT CATTTATTCA ACAAATATTT ATTCATAACC
75151 AATTAAGTGC CAGATGATGC ACATATAGAC TTCTTGTCTT GTTGTTGCAT
75201 TGCATATTCC ATATTTCAGC TATCCTGAAT TGTTTCAAT TATTCATAAG
75251 TTCTTTATGA ATTGTGTTCA TTCCATTTGG AATATTCTAC CTTGTTTGAT
75301 CAGCATAAAG ACTTTTCGAG ACACTGCAGC AGCAGTGAAC CTAAATATGT
75351 TTCCTTGACC CCTACATTGA ATGACACCCC CTGTGATATG TTTCTGGAAG
75401 CAGCAATACT TCCCTTCTTA AAATTACATT ATACTTTGGG GCTTTTATTT
75451 AAGGTATGTC TTTCTGATT TACAATAGTA GAGCTTGTTT TTTCACCCTT
75501 TTGAAAGACA TCAAGATGCC CATGATGATG TCTTGCATGT AACAGGGGTT
75551 TATTTGAATT TTTAAAAGAA GAATAAAGTA ATTTTAAAT GAATTTCAAT
75601 TTAAATTTTA GGAACAAT TATATAAGT GAGATATGCT TAAATTGAAG
75651 GACAAAGTAG TTCTGTAGGG GCTACTTCTT TCAAGACTTT AGCAACTTTC
75701 CATGTGGGGG AGTGATTAT GTGATGCATG GAAAATTACT GCATATTTAA
75751 AGCTTATCTT AGAGCTATAA TAAAGCAGCT TATGTTCTAA ATCTTCATGT
75801 CGTAAATAGG TCCAGAAGGG ATTTAAAAG CCTTAATCCT TACTTTAACA
75851 CAGCACAAGT CACTGAAGTG AAACCTGCTG AAAGGATTCC TTTTATGTTA
75901 GGCAACAGGT AGCTGAATAT ATCTACAGAA ATTGAAAAAT TGGAAATCTT
75951 TTGCTCAGAA ATGTGGGAGG GGTGGAGCTT AAGGTAAAAA ATAACAGTTA
76001 ATATCTAAAT TGATCAAGAA ATATGAAAAA ATAATTTGCT AGGTTTAAA
76051 ACTAACAAAA ACCATGGTTA TAAAGGTTTG AATATATATA GGATAGTTAG
76101 ATTGTATTTT TGTAATATTA AAACCTCAGCA TTAAATTTAA TGAACACAAA
76151 GTGATTCTTA TCACATTGAC CATTGACATT ACATGGAAAA AATAGTCAGT
76201 TGGACTAATT ATGTGTCTTT CCATGGGTTA TTAAGGTAAT TGTATGGCAT
76251 ATAAATTTAT ACTGGAAATC ACATTGAAAT TCACTTTTAG AGGCCCTTAA

Fig. 2 (cont'd 42)

48/ 124

76301 AATATTTCTG TAATATATAT TTTTAACATA TGATCTTAAA AGATATATTTT
76351 GGAATGACAC AACAGTTTTA TAGACAGGCC TGACTATCAC ACAACCACAC
76401 ACCAATTTGT GAATGTGTTT CTATTTCCCTC TAAATTAATG CATCACATTC
76451 ATTAACAAAG TTTGATAAAT GACTATAGTC TATAATAAAA TATTTTGTGTT
76501 TACAAACATA TTTAAACACC TGCTATTAAAG TATAGGCATT ATCAGATCTT
76551 AAAATACAAA GATTTAAAAA ATTACCCTGT GGTCAATGGAG CTCACAATCC
76601 ACTGCAAAAA TAATGTTTGT GATAAGAAAT TTGAAAGTTG AAGGTAATAG
76651 AAAATTTTAC CTTTATTTT CAAAATGTAC CATTGCTTTC TAAGTCACTA
76701 CTTCGTGTA AATATGGAAT TGTTTTTCCT TAAGATATAC CAAATATAGT
76751 TGGATAACGC ATGTATTAAA ATTCTGTCAG CACTAAGTTG TTTTTTAGAC
76801 ATAGTGATAG GCAAACATAG TTATATTGAA TGAAAAATTA GAATCAAAT
76851 TATTAAACAC TGTGTACTGA TTGATACCAC ATGCCATATG CTTGTATAGC
76901 AATACAAGGT TTGGAATTTA TAATGGTAAA CAAAATAGAT ACGGTCTTTG
76951 TCTCCATAGA ACTTTTAGTC TAGTGGGAGA GCAGAAGGTA AAGGAATGTA
77001 TGTGATCATT GGTGAAGCTG AACATGTATA CCCAAACAGT TATAAGTTCC
77051 AAGATGGACA ATAATGGGTG CCATAGGGAA GGAGGGTACC AAGGAACCTA
77101 CTGGAGGTTA CATAGGGAAG ATTATTCCAA GGTAAGTAATA TTAAAGTGAA
77151 TATCCAAGGA ATAATTGTCA ATCACTTTAT AAGTACTGAG GGAGGAGTAT
77201 TTCAAAAGAG CTTTGAGGCG GAAAATAAAT TAGTTCCTTT ATGGAACATA
77251 TGTAAGGAAA ATACTAAGCA AACATGTAAT AAGAAGAACA CGGTTGATGA
77301 GTTAAGAAGT GACAAGATTA CTGAAGGATT GTAGGCCATA TTTAGAAGTT
77351 GGATTTTTTA TCTATTCTTA TTAAAGTGAG AAGTTATTGA AAGGTCTTAA
77401 GTGGGGGAGT GATGATGAAG TTTGCCTTTT AAAAAAGATT TTTCTAGCTA
77451 TTGTTTATAG AATGGTTTGA AGATGAATAA GTCCAATAGC TATACTTGCT
77501 GTAAAGGTTA TGTTGGTAGC TTGAAGTGGG GCAGTGGTGA CACAGAGGAT
77551 GGGAGATGGA AAATGACGAG TGAACAAACA CATACCTGAA AATTTAAGTT
77601 TAAAAATAGA CCTCTCCATT AATTCAGATT GCTGATATTC ATTCGGTTAG
77651 CCATTCTTTA CTGAACCTTA TGATGCCCCA TATACTGAAT TAAATACTTA
77701 CAAGCACTAA AAAAGAAATT GTTAGGGAAC AGTAAATGC ATTCCTTCA
77751 TTTCACAATA TTATTAATAT TATGGCTTTG CTAATCTTTA TTGGTGAATG
77801 CAGTCATAAT TGAAGGTAAC TGATACTTCC AAGGACTACT TTTGACCTAG
77851 GATTACTATC TTTTAAAAA TTTAGTATTA AAGAAGTCAA ACACAATTTA
77901 TTAATTCCTG ATATAATAAA AATTCGTAAA TACTTTAATA CTTTGTGCTT
77951 TTCTATTTGT GAAAGTTAAT TATTAGGAAC GAGCTAGCAA ATGCTACTTC
78001 TTTTCAAAA AGCTAATGGC CAATCACAGC AAAAAATTA AGCACTAAGA

Fig. 2 (cont'd 43)

49/124

78051 AATACCTACA CATATTCTTC TATTGCCCAT TTATATGACT TCCATAATAG
78101 TTGATTAAAG GATACCGGAT TCCTTTATTG TTGAATTAAA ACCTCCTACA
78151 TGA AACCTT GATTTAGGTT TAGAAGTTGG TAATGTTTTG GCATGCAAAA
78201 CCAGTTAATG TTCTCATCAT TACTTTTTTAA AACAATGTGA AGAGATGAAT
78251 TCTAGGGATT ATAAAAAAA AAAAGCTGTA TGTGTTTCTT CCTATAAAAT
78301 TTTTCAGCAT GATTGCCTCA GTAGAAAAAT TAAGGGACTT ATTGATATAT
78351 ATGTATATGA AGGTGAGGAT ACACATATAC ACACACACAT ATATATGTAG
78401 GTAAATACAT ATATTACATG TCTATCAATC CATACATACT CATTATTAT
78451 ACGTTTTGAA AGCAACCAGT TATAGTTTTG TGCCATGGA TCATTTTTAC
78501 TATTCAGTAA ATCAGTCAAT TGAAGAGGCT TGATTTTATG GTATTAGTTT
78551 TTTGGAAACT GTCAGCTTTA TAGTAAATTT TGACATCTTA CAACTCCAC
78601 TGAGATTTTT TTGCTTGACT AATCTGCCTT GATGCCAATA AGTATATTAA
78651 CGGAAATGGA CTAAGCAA ATGTGACTTG AAGCACAATT TTGTAAATTT
78701 TCTTAGTGTC TCAGTAATAC TTAATACTAG TGCATTTTAG GTAGGAAAAT
78751 TTTTCAGTTG TTTTATTTTA AATAACTATA AATCTTATAG TTGCTTGAT
78801 AAAAGAAACA GATACCTTTA ACATGATTAA ATATCAAATG CTATTCTCTT
78851 CAAATATCT TAACTAAAGA AGCACTGCCT GCTCTTAGAA GTTAAGCAAG
78901 GCCATACCAT ATGCTGCGTA CATGGCTTTT AACACAATGG ATATTAGAAA
78951 CAGCCTAAGG CTGAGCCTGG CTCCACTATT TTTTCAGCTAT GTGACCATGT
79001 GAAAGTTACA TTTAGTAATT AAACCTCATT CAGTAGTTTG CTTTAAGAAT
79051 AAAATTAGGT ACTCCGGGG CATATCAAGC ATATTGTAAA ACCTAGTTTG
79101 ATTATTATTT GTTATTGGTA TTACTATTAC TATTCTATAA TAAGTCATGG
79151 GCAGGCAGTA GGGGTACATT GGAAGAATTG CACTGTCTTA AATATGTCCT
79201 CTGTTTAACT CACAACTCA GTCTACCTAG GCTTTCTTTG GAGGATCTGC
79251 CTTTCATTGG CTGTTTGACT TTGGCCAAGT TACTTAACTT CTTTTCATT
79301 CAGTTCCCTC ATCTGTGAGA TTATGTGCTT ACATGACTTC AGGTTTTGTT
79351 TTGGCTCTAA TATGGTATGA TTCTATGAAA TGGAAAGTTA ATACATTGG
79401 CTCTAGTAAC TGTATTTGAA GCACAAATAT TAAAAAGCAC AATTAATTCT
79451 CATCTGAGT TTCCATTAC TCTTTTAAAT TAATCATTCA GAATAAATCA
79501 TTTTGGAAGA GCTGCTTGAT CCAGGTATTC AGTAGAAATC ACTAGCATAG
79551 CATTTAATTT TAGACAAAAC TGAGAACTCA TTAACTGCC AGGCTATGG
79601 ACTTATATGA GATTCTCATT AAATCTTAAT GTAGATAACT CAGTTAATTA
79651 AAACAAATAT GGTGTACTT TATTAACTT CTAAAGTCAA AACTGCATTG
79701 AAATTATCTG TACAAAGCCT TGTTGACCTT TATTAGAGAA CTGCCCTCTCA
79751 AAGACCTAA AAGACTTATT TGTTGAGATC GAGACTCTTC ATGAGCCAAT
79801 GTGATACTCT CCCTCTATTG CTAGATCTTC GCATCAGAAG ACAGCATTC

Fig. 2 (cont'd 44)

50/124

79851 TCTGAAAGTG TTTCTAGTAT CAACAGTGCC ACAAGCCATT CCAGTATTGG
79901 CAGTGGTAAT GATGCCGACT CCAAGAAGAA GAAAAAGAAA AACTGGGTAA
79951 GTTACCATCC TTCATCTAAT TCAGAAGCTT ATTAATGCAT AATGTGTTAG
80001 GCCTTTTCTT TTGGGGCTTT AGTGATCTGC AGTAGTTTAC AAAGGGTCCC
80051 ATTCAAGCTA CTGAGACCTC AAATGCTGCA CTCATCACCA AAATGGAGT
80101 GGCATGTACT GAAAAGCATA CATTTTAATG TTGGGACTAA ACTTGGGTTT
80151 GAATCACCAC TATATCTAGA CCTTTTGAGG GGCCTGAATT TTCTAACCAA
80201 TAAAAAGACA GTTAATAGCA ACTATATTTA TTTGTGAATA TCATTTATTC
80251 ACAGATGTTA TCTAATTTTT CTATAGTATA ACTATACAAA CTATGTAGTA
80301 TAACTATAGA GTTATACTAA AGAAAAATAA GATAACATCT GTGAATAAAT
80351 GGCTTAAAT AGGGGTTTAT TGTGGGCATA GAGATGAAGG AAAAGTGAAA
80401 AAATGATGAT GATGGTGATG ATGATGGTGA TAGTGGTCTT GGAGGAAAAG
80451 GAGAATGGGA GTTAATAAAG GGAAAGAATA AACAATGAAA CTCTCATTCC
80501 ACCTTTGGAA TCGACAGGGC TTACCGTGTG AATAGTTTCA CCCTAAAAGA
80551 AATCAACEAC ATTAGTGCTT GCTTGATGTT TTTAACCAAG AGAATATAGC
80601 AGAAATATAG AAATGCACTT TAACAGAACT GTACCTTAAG TTTGCTAGTG
80651 ATATAATTTA TGATATTGAT CAATAGCTAA ATAGCCCAGG GGAAGATACT
80701 GTTACTGCGA AAAATTTAAA AACAATGGAG TCAATGATTT CTTTAAATAC
80751 CAAAAAATAA ATGTAGATTT TGAGTAAATA CAACTCTTGA TGAAATCCAG
80801 ACATAATTAT CAGAGGATTT TACTGGAGTG CTTTCTACAA ATAATGAAAG
80851 AAATATCTTT TTATCTTAAA AAATGTTTAT ACAGGTAATA TTTTAAATA
80901 CTGATCAGCC TTCATTCCCT TGATTTGTAA TTCCCACTC TTTTATGTTT
80951 CTGCAAGGTG AACTCTAGAG GAAGTGAGGT GAAXATAAAC CGTGGACAAT
81001 TTGGCATGGA TTTATAAAAA AACCCTACCT TGGCATGAAT GCTATCCATT
81051 TTGGCAGTAG GCTTTTATAC CTTTAAAAAC AGATTACCTT GTATGTCTTT
81101 TCTTTGTGTC TTTTCATTTT AATCTCAAAT TTAAAGAGA TGTAACCAA
81151 CTTTCTGAAT AGAGCTGTAG GGGATACCAA TTCTGGTTTT GAGTAGTCTG
81201 GGGTTGAAA ATTTGAATAG AAAAATCACA ATTAATGAAG TGTTAGGTGA
81251 ATTTGATTTT ATTTTGCTTT TTAAGTTTGT ACTGTCAGCA GGACATGACT
81301 TGATTGTAGC GCTAAAGTGG CCATTTAAAA CAAATGCCT TGAAGAGAGA
81351 AGCATTGGGA ATGGAGATC

Fig. 2 (cont'd 45)

51/124

Human genomic sequence

1 GAATTCCTGG TGGAGAACAG CACATGTACA GATGGGGTGA GAACAGCATA
51 CGTACAGGTA GGGGTAAGCT GGTGCTATAT GAGAAAGCAT GGAATAAGTT
101 ATTAAGTTTG ACCTGCTTGG GAACTGAGGG GCAGGTGTGA GGGATGAAGC
151 AGGAGTAGGT AGGGGCTAGA TCACAAAAGA TCTATGCCAG TGTTCCTCAC
201 AGTGTGATTC CCAGCCCAGT AGCATGATAT CACTTGGGAT CTTGTTAGAA
251 ATACAAATTC TTATACATCA CCCTGGACTA GACCACCTGA ATAAGAAAAG
301 TTGGGCATGA GGCCTACAAA TTTTAAAAA AGTCATACAG GTGATTGCAA
351 TGCATGCTAA AGTTTGAGAA ACACTCTTTG CTGTGGTTTG AATATTTGTG
401 TCCTTCCAAA ATTCATGTAG AAACCATCTC CAATGTTATA GTATTAAGAG
451 GAGGGACCTT TGGGAGCTGA TCAGATCATG AAGTCTCCTT TCTTATAAAG
501 GGGATTAAAA GCCTTGGCCC TTTTACCCTT TGTCCATGTA AGGACACAGT
551 GTTGAAGCA GGGACTGGGT TCTCACCAGA AACAGAACCT GCCAGCCTCT
601 TGGTCTTGGA CTTCTCAGCC TCCACAATTG TGAGAAATAA GTTCTCTGTG
651 TTTATAAGTT AACCAGTCTC AGGTATTTTG TAATGGCAGC ACAAAGGGGC
701 TAAGAACTG TTCTATGCCC TAACAAGAAA TGTGGTCACT TTCTGAAGG
751 AAATGGGGAT ATATATAAAG ATGTTATATA AGACTCGTAA TATTTATTTG
801 GAAGGCTTGC TCTGCAAGCA AGGTGGAAGA GCAACATGAA GGAAGCGTGG
851 TGGAGGTGAG AGGACTGGAG GTTAAGTTGG TAGGGAGATA CAGGAAAGAA
901 GCTTATGACA CTTGAGTTAA AATGTAGCAT CCTTCCTATG TGTAGGGCTC
951 ATAAAAATGT ATAGTCTAAG ATAGAACACA GAATACTCTA TGAATCCTGC
1001 CCACAAGGTG TTGGTAATCT AGATTCACTT TTTTCTCTG ATAATGCCAT
1051 CCATATGTAT GGAGCGTCTA CTACTGTATG CCAGAGTGAC TCTGGAATCG
1101 GTTTGGTTGA TCTAGACAAG ACCATAAGGA GAGTCCCCTT ACTACCTCTT
1151 CTCCAGGGGA GGGATTCAAG TTGAACTAGT ACTTCAGAGA CTGTTTAGTA
1201 ATATCATGCA TGAAAGGTGA TGGTTAGGAC AGAAAAATAA ATGGATTGCA
1251 TCATAATTCC TCAGGTCTCT CAAATATGTG GTGGTCTCAA ACCATGTGAA
1301 TTGGTCTGCA CATCCTGTTT GGGTTGCGTG TCAGCAGTTG AGATCTGAGC
1351 CTTATTTGTA ACAGTGAAAC AGTGAGAGAC CTGCCCTTCA AGAGCTGTTT
1401 TTCAGCTAGG AATAGAAAAG GGCCAGGCTA GACTCCTCTT TCTGCTGGAT
1451 CTTGCTTCTT CTCAGCAATA GAAGTAGACC TGCCCTCCTA GCTGTAGAGA
1501 AAAGGTGCCG GTAGGCGGGC AGGTGAGCCT GTGGATAATC CTGGAGTAAA
1551 GGTTCAATAG ACCTTCAAGT CTATCCTACA GGATTCGGAG TGAGGGGAGA
1601 GAAAAGGAGA CGCTTCTCTG GCTGAGAGAG GAAGAGAAAA AAAAATCCCA
1651 GATATCTGAC AGCTATATCT TCCCATCACC ACCTTCCTCT AAACCCATGC
1701 CTCTCTGTTT AGTAGGACAT AAAATGAAGA GTGACCCACC CCCCACCCCC

Fig. 3

52/124

1751 AGCCCATCCC CCGTTTGTAG GTGTGCTTTC AATGAAAATA AGTCGGTGTT
 1801 CATGGACGGA AACTAGAGCA GCTGAAAATA GATGCAAGAC TTGTTGAGCA
 1851 TACAAATCAT TTCCCCCTTA GTCTCCAAGG GAGGAAAAAA AATCCCTCTT
 1901 ACTCTCCTTG CAGCCTGTGT TCTGCATTCT GGAGAGGAAG CTGAGGCTGG
 1951 TCCTCAGGCG CTCCTCCCGC CGTTCCCGCA GGAAACTTTT CTCGCAGGGC
 2001 CCGCTCCGTC CATCCCGCGC GGTTCACAAG CGGTGGGCCT CCCGTGGGCT
 2051 CCTCTCCTGG GCAAGGGCCC AGACCCCGCG ACGCGCCTGT CTCTTTAAAT
 2101 TCCAGCTGCG CGGCTGGGAA ACAGCGCCAC TCGCCGCCCA GGCCGGCTGG
 2151 AGGCTGAAGA GCGAGCTCGC GCTTTCGCTC CCGGCTGCGC GCCGCGGAGA
 2201 GCTGGGCTCG GCCCGCGGGC TGCTAGGTGG CGGCGGCGCG GGGCGGGGAG
 2251 GCGCGGCCCC GCGGAGGAGG GAAGAAAGAG CGAGCCGGGC CGGGAGAGGC
 2301 GCCGCGCCCC GTCCCGCGCC CGGTCCCGCA CCCGCTCTCA GCGGCCCAAG
 2351 CAGTTTCTTT CTGGGTGACA AGAATGTGCC TCGGTTGGTT TTTCTTTTTT
 2401 TTCTCCATCT CCTTAAGACG ATTTCCATAG TAACCTGATC AAGTGGCTCA
 2451 AAATCGCAAA CCTGAGGATT TCCGCGGCCC GCCGCAAGA CCTCGGCCAG
 2501 GTAACGCTGC GATCTCCTCC TCTTCCATTG CAAACCGCTG CGCTCCTTGC
 2551 AAAGTTCCTT TTGTGAAAAA TCGCCCAGCC CAAGGGAGCC CGGGGTATTT
 2601 GCAACAGCGT GTTCATTTCG AGGTGCCTGT CACGGGTCTC CTCCCTGCTG
 2651 CTTCTCCAGG ACCCATGATG AGATTATTTT TAAAAATTGT TTTTGGTCGT
 2701 CTCCCCCGCC CCTTCCCTTT CTTTATTTTT TTCTCTTCG CTGCACTCTT
 2751 CTCGGCTTTT CCCCTGACAC TACTGATGGG GGTGCGGGGG GACGTCGGGG
 2801 ATGGGGGTGG CCAGCGCGGT CCTGGGAGTG GCGGGTTCGG ATGGGCTGGC
 2851 TGCGGTGGGC CACTTTGGGC ATCTCGGCGT GGCTGCGCC GGGGTACGG
 2901 GGAGGGCTGT CAGCGCCAGG GCGGCGGAAC CCGAGGTCTC CAGACGAGTG
 2951 AGGGAGGGAT GCAGGCTTGG GGGTGATGGA GCGCTTGGCT GGTGGCTGGT
 3001 GAGCGTCCAT ACATCATAGC TCTCCTTCCC ACTCCCCCGC CCCTCTTCGG
 3051 GATTCTCTCT TTCTCTTTCC CCGTCCCTCAT TTCTTTCTTC CTTTACTCAC
 3101 CACTCGCTTC ATTCTCTTCC TTCCATTTC TCTTTTTTTC TCCCCTCATT
 3151 TCCTTTTTTT CTTTCCCTTT TTAAAGAAAG GGAATCGTT TGTAACCCTT
 3201 TCGTTCTACC AACGTGGAAT AGCTGTGAAA CCTGCAGCGT GGTCACCTCA
 3251 GCCTGGTCGT TTTCAGACCC GTCCTCATCC ATCAACATAT TTGTTTCCCG
 3301 AGTCTATTGA TCTCCCTGAA TTCTACAGAA ATGCATTCTA AGCTAGGCGC
 3351 CTGTATGTCA GAATCAGTTC TGCAGGTAGC TTCCGTGCTC CAAGTATGAC
 3401 ATGTATTGTA AGGGCTGCAT CTGTTTTAAA CCCACATAAG CCATGGGTAT
 3451 AAATAAATGT AGCTTTGAAA AAAAATCTGG CTTATTCTA GATAAACTTC

Fig. 3 (cont'd 1)

53/ 124

3501 CCTCTTAAAT TACTGATATA CTCTTCTCCC TCTTTGACAT TTAATTTTAG
3551 GAAAGTTGGG AGACAGGTTC TTGTCCTCCA GTTTTAAAGG AGCAGGCAAC
3601 TTCTATTATC TTAATTTTCT CGTCTTTGAA CATCACTCAC GTTGCACATA
3651 CCCAGTCAGT GGAACGAGTG GGTCATAATT AA

Fig. 3 (cont'd 2)

Human genomic sequence

1 CCTGCATTAT TGTTTTTATC TGACTTCCAA TTTTGGTGTT CCCTGGGTGG
51 GTGGGTTTTTC CTGACACATT TACAAGATGC TTTTGGCAGG TTGGCTGGAA
101 TTTGAAGGCA CATTTAATTG TAGGTGCAAT AAAATATTCA TTTTCTCTTG
151 TTCTTGTTTT GAGATGTCAT GCCCTTTTGG TCACTTATAT TTTGGTGTGA
201 CTGTGTGTGT GTGTGTATGT GTTTGTGTGA AGGATTTAAC AAAGTCTGTT
251 CTAAGTGTCA TGTGATTTGA AGTTAAAAGG TATGTTAGTG ACAAGCCACA
301 AATTTCTCTT ATTTATAGTA CATTGATCCT GAAACCATTT TTTCCCTTGT
351 GATTTCTTCT GTGCATGGAT CATTTAACGA AAGGTTGGCA ATGATGAGCT
401 ATTTTTTTAT AATAGGAAAA AAATTCCCTCA AGTTTACTTA CCAAGTCATA
451 TTTTTATACA GAGGGATTAG CAAATATTTT TGATCTAATA TTTTAATAGA
501 CTGAATTGCT GACCACTGCT AATTACCAAG AATATATTTT CTTAATTCTG
551 AAATTGCTGT ACCTCTCAAG TTGTCTGGAG GACTCCAAGT GACCCAACCT
601 GTAAGTCATG GCAACAGGAA GTGGTTGTTC TGGGTGCAAG CTGAAGTGTG
651 CACATGGACC CGTACTTTGT TAGCACTCGG GGACTTGATA TGGAAAGAAT
701 TAATGTACTG GCTTTTTTGT ATAGATGAAT GTTAACCTTC TGACATTAGT
751 CAGAACTACA TCTCCCAAGC CTTGTTTTGC AGTGTCTGTC CCTTTGCTCT
801 TCACTTACAG TAAGTCCTTA CTTAACTGAC TTGATAGGTT CTTGGAAACT
851 GCAACTTTAA GCAAAAGGAA GTATAATGAA ACACTTTTAT CACAGGCTAA
901 TTGGTAGAAA CAAGACTTAA GTTCCCATGG CATATTTCTG GTCACAAAAA
951 CATTTCCAAA CTTCTCAAAA CACTTCAATA TTAAGCATT CAAATACATGT
1001 AAACATATGA TATATGTAAG AAAGGTTACT ATAAACCAGA TCAATATTTA
1051 CCCAATTATT TAAGTTCAGG GTCTTAGGTG GCTGGAGCCT ATCCGAGTAG
1101 CTCAGGGCAC AAGGCGGGAA CCAGCCCTAG ACAGGACACC ATCCTGTTGC
1151 AGGGCACGTT CACACATGCC CACACGCAGG CTGGGACCAT TTACATGTGC
1201 CAATTCACCT ACCATGCACA TCTTTGAGAC GTGGCAGGAA GCAAGAGTAC
1251 CTGGAGAAAA TCCATACAGA TATGGGGAGA ATGTACAAAC TCCACCCAGA
1301 CAGTGGACCC AGCCAGGAAT CAACATTTGG GCAACATTAT AATGAAACGA
1351 AGTTGAATGA AATGATGTCG TTCCACGACC TGCTGTACTT GAGGGGTGTT
1401 ATAAAATTCT CAGAAGACAG AGGTTTAATG CTATCTTTTT AATAGAAAAT
1451 AACTTATAGA GAAGTGTGCA CATGTGACTT TGTGTGTAGC AGGAATCAT
1501 AGGATGAGAA TCAGACGTAA GAGGTGGTGC CAACATGAGG AATGTTGAGA
1551 TTCAGGGAGC TGTGGATGGA AGTAGAAGCC AGAAGGCCAG GGTAGGTTT
1601 CTAATTCTTA CTGTTTCAGT TATTGCAGTG TTGGCCTGTT TATTCACAGA
1651 TGTCACCTAG CTTTGTTCCT TCAAGAAGAA AAATGAGCAT AATCTTTCCT
1701 GTTATGAATT CTTAAACACA CAGGACATAA CCACAGACAC AGAGGTGCAC

Fig. 4

1751 ATATGTAGCA GTAATGGATA CTAAATGATA CACTCGGAGG AAACAGAAAA
1801 GACTTCTGAA TAGAGACTGG AGATACTTCC TTGGACCATT GATGAATGGG
1851 CAATGATGCA TTTTGTCTT CCATTCAGAA GGCTAATATA TTGCTCTCTA
1901 TGTTCATATG ATAAAGGCAG TATATGCTCA AGGATGAATC ACATAATATG
1951 CATAATAAAT CCAGCAAGCA TTACCCCTTTT ACTTATGTGA CTGCAAGTAG
2001 GAATACATTT CCCCCACTCT AACCATGTAA GATTCTTTTC CCTTCTCCCA
2051 TTTTGTAAGC AAAAGTAAGT TCCTGAAAGG TTAAATGGAC CTCAGGATGG
2101 GAAAAATCCC CAGAGCTATC TTTCTGCACA GACTTCATTT TTTCTCCCAA
2151 GTCTGACTGT CAACTGCGAT ATCTGATATG AGGCTCTGGT GCTGATGTTT
2201 CCATAGGTCA TCATCCTTCG GTGTCCCAGA TGAAGTCTCA GGTCGAACAT
2251 TGCAATAGCA CAGATTCTGA ATTTAATGCA TCATTAAAGT TGGTTATGTA
2301 ACCCAATGGC CTGTGTTAAAC TCCAGATTTT TAAAATTATA TGTATTTACT
2351 ATTCTCTTAT TTTAGAATGA TCTCACAATG TTCACAAGAA ATAAGCCCAG
2401 TCCCTGCAAA GACTTTAAAA GCTGCTTGTT CACATCATTA GATTGTACAA
2451 CGCTTGATCA ATGACACTTT TTGCTAATCT ATGCAACATT TTTGTAACAA
2501 TTGTGCACAT TTTAACTACT TCAGATAATC AGGACCTAGA GACTTCAAGA
2551 TCTGGAAGCA TTGCTGGTGA CATAGAGCAA AAACCTTCTT GAGAATAGGA
2601 AGTCAGTGTT TTGACAAGTG ATTTATAACA GTTCAGGTAT AGCCAGGAAG
2651 GTTTGAAACA AACCTTAAGT ATTATTTCTT TCATCTTGAT TAGTATATAT
2701 TTATATGTGA TCTATTTATG TATATTAATA GATTTTGGG TCTTATAGCC
2751 AGCTTTCATT TTTCTCTATT GGAAAAGATC TAAGTCCCCA TCCTTCCTTG
2801 GTGGCTTTTG GTAGGTTTGT AGACAAAACA TTGAAGAATC AATGGTACCT
2851 TTTATACATT AATACTGCCA ATATGACCAT AAAATCATAT TTTTGGGAA
2901 TTTATTTCCC CGATCAAAAG AAGCATTTGT TATTGAACAC AGTCTTATGC
2951 TACCTTATTA AGATGTATCA AACACCCTGA TTGATCAAAA ACACCTCAGT
3001 CCATTTTAAG GCAGTATTGC CCAGCAATTA AAGATGTAGC TTCTGGAGGA
3051 GTCTTTCTGA GTTTGAATTC AGTACTCTTC CACGTACTAT ATAGGTGATC
3101 TTGGGTAAAC TTCTTGAGTC TCAGTATCCC CATCTGTAAA ATTGTTGTAG
3151 AGAAGAATTT TTGTGATGAT TAGGTGAGAG AATATATTAA TGTAATATTT
3201 AGGAGAGCAA CCAGCATGTA GCATATATTC ATTACATATC AATTTCTATA
3251 TTATTGATGT TCATACTGCT GATGTTGAAA TGCACAGGAA GGCCACAGTT
3301 ATTTTCTGTT TAGATTGATT TTTCTTTTAA AGTCTGAACA TAAACTGTAA
3351 TACTGTGCTT ATTTATGTAG GAACTGTGAT CTCGTCTCCT CCTTTTCCCA
3401 TCTCCCCCTC TCTACCTTAG TTTTCTCTTA TAGTCTCAAG CTGAAAACAA
3451 TGACCAGGTG CCTAAGAGAT AAGAATACTC TTTCTTTTGA ACTCATGGCA

Fig. 4 (cont'd 1)

3501 TTAGCAGTGA CCTGGATGAG ATTGGAGGCT ATTATTCTAA GTGAAATAGC
3551 TCAGGAATGG AAAACCAAGC ATTGTATGTT CTTACTTATA AGTGGGAGCT
3601 AAGCTATGAG GATACAAAGG CATAAGAATG ACACAACAGA CTTTGGAGAC
3651 TTGGGGAAAG GGTGGGAAGG GGGTGAGGGA TAAAAGACTA CAAATAGGGT
3701 GCAGTGTATA CTGCTTGCGT GGTGGGTGCA CCAAAATCTC ACAAATCACC
3751 ACCAAAGAAC TTACTCATGT AACCAAACAC CACCTGTTCC CCAGTAACCT
3801 ATGGATATAA AAAAATTAAA AAAAAGAAAA AAAGAAAACCT CTTTTTTGCA
3851 GGGGGCAGGT AAAGGGTAAG AGGGCATCCC ATTTTGTGAGT TTCTAGAAAA
3901 GCTT

Fig. 4 (cont'd 2)

57/ 124

Human genomic sequence

1 CTGCAGGAAG CAGCAGCAAG GTCCAGGGAG CCTCTAATTT AAATAGGAGA
51 AGTCAGAGCT TTAACAGCAT TGACAAAAAC AAGCCTCCAA ATTATGCAAA
101 TGGAAACGAA AAAGGTAAGT GTTTGTTACA TCATTATGAC ACAAGTCCAA
151 CATGAGTCTT GTGAATTGCA TGCTAAATCT AATATTTGAG CAGCGTAACA
201 ACTTTGGGCC TAGAGATGTT ATCAGTGGAG TTTCTTTATG TTTCTTAAT
251 GTCCCTCCT GACTGCCAGC TTTCTTATCT GAAGAACATT TTAAACAAAT
301 AAACATCTC ATTTTAAAGT AGTTAGTTAT ATATGCAAGT ACAAATACTG
351 TTTCTCAAAA ACAGGTCCTT CCAAATGCAT GTAAATCACA TTTCTTTATG
401 TCTTTTTATG TTTTGGAAAA TGTATCCTGA AATCATAAAG CCATATTGAA
451 TTTATCTGAA TCCTTAACTT CAGTTAAGGT AAGAGCCATA AGTGTTTTTG
501 ACAATTAAGG TTGGAGCATC AAAATTTGAA ACATAATTAC AGTAGGTTTT
551 TATCTTTGCA AGCAGCAGAT CCCAGAGATA TTATGACCTC AGTTTTCCCC
601 AAAAGACAAA TTATTCATAT TTGTTTTGTT TTCTTGAATT AGTGCATAAT
651 ATAAATATCA AATCACAAAA TCAAGGACAT TAAATGAAAG TGTCTGTTAA
701 AGGCATATTA TAAATGAATC ATAAGCCACA CAGTTCTCTG TGATGTACGA
751 AGTGGGCATT TAAAGAGGTG CTGATTTGAT GCTTGTCACT GAGTAGCAGA
801 GAGGACGGGG ATGAGTATGT GTAGTTTACA CCTCAATCAT GAGGAAGTGA
851 AGAACTTGTG CTGTTATAAG TAGTATGGCT GTGTGAGGAA CTAGGGTGTT
901 CTGCTGGATT TTGAGGAAGT ATTTTCAAAT CAATAGAACT TCAAACTTTT
951 CTTCAGAGTG TTGGGCTCTA CATGGAAAAA CACATGAAAT TAAAAAGTGG
1001 CACAAATGTT TAGTTAGTAG AACATCTGGC TAATTGGGAT CAAATAATTC
1051 AACCATGTGG GAACGTTTTT GCTCAAAATA GATAATTGTG AATTGTTTCA
1101 TATAGGCAAA TGATTAGACA ACTTCCTCTT CCTCAAATGT GAACGGACAG
1151 ATGTGATCTA GAAGCAAGAC ACTCTTTTGT GTAAATATTC CCTTTGGCCT
1201 AAAGCAAAAG TGGACAGACT TTAAACACCT GAGAGCAGAG CAGTGTGTGT
1251 TAAGATTGCA ATATCTTAAG CTCTTGAGTT AAATGGAAAA TGAAAAACAA
1301 AAGTGATAT TTGGAAGTTA GGAATGTTTT CTTTAAATA TAAAAATAAA
1351 TTTTAGATTT AAGATCACAA GAAATATTAC TGAAGACTTA TACTCTTCCT
1401 GGGGCTAAGG GAGGTGACAG TCGCTCATCA GAAAAAATAA AATGCCCTCA
1451 TTTCTTAAT TTTCTAAAAA ATATAATACA AGTTCAGGCT AATACTTCCT
1501 GTATATGTGG GAAATTTCTA GGGGAAGCTA ACAGGCTTAG AAATAAAGAT
1551 GTGTTAAATA GACTACCAA GTGTCCAATT AAGCAACACG ATACCACCGT
1601 TATTGATATT CTAGCAAGAA ATTACTAGCA ATGTTTGTA ATAGACTTAG
1651 AAATGCATTT GATGAATTAA CACTTTTATA TCTTAATTTA TCTGAATTTT
1701 TCTGTAATGT GAAAATGTTT TATTTAACTT ATTTCTGGCA TCTATTAGTA

Fig. 5

58/124

1751 AAATTCTGAT GATATACAAG CATTAATATT TTTCCATGGC CACTCAATTC
1801 ATACATACCT TCCCTATCTA TGCTTAGAAG GCAGTGCAAA ATTAGATAGT
1851 AGCAATATTG ATTATAACCA CAAGGTGGAG ACAGATGTCA TGTAATATGC
1901 AGTCTGCTCA TATAAAGCAC ATTTTCTTAG ACAAGAGTTT TCATACGATA
1951 TAATAAAGAC ATCTGGAATT TGTCTTGAT GCAATATGAA ATTTGCTATT
2001 AAACGTGGAG TTAAACTTT ATGTCAATAG ATCCAATAAC AATGTTTATA
2051 AATTAATCAT TATGTCATGC TGTATTTCCA AAATACTATC TTAAATTATA
2101 AGAGCAAACG AGGTAATAA

Fig. 5 (cont'd)

59/ 124

Human genomic sequence

1 GTACATTTTT TAATAAGAT GTTTGTTTTA ACTTTTGA TATGAAGATT
51 TCTAGTTCTA GAATAATGTT TATAAAAATA TACAAATCCA TCTGGTGATG
101 AGTTGACCTC TATCACAAC AGTTTGCATA TATAACTTGG GTGTGACCAA
151 GCAAGGTGAG AGTTAAGAAC TTTTAAACT TACTGTATTA TATTGATAGA
201 ACTCAGAAAG TACTAACTTG AATATTATTA TTCTAATTGC TTTTCCCTTT
251 TAGTTATTAA AAATAAGAAT ACTTAAATTA ATAACAAGAT CTTTACTGG
301 CAGGATTAAC CAAATTATCT GTAATGTGTT CCTCGAATGC TTTTAAGTGG
351 AAATATACTT TATACATTCT TTAACAAC TCAGAGGATG AGTTACATAA
401 ATCAGTTCAG GAATCTATAG AATCTGTAAT ACATAGTAAA GGTTTATTCA
451 CAATTAAAC AATTTCACTT CTATATTAAA AAAACAAATT GTTGAAAGTA
501 CAGTGGCTTT TCATATGTAT GATTTGTAAA ACAAATTAGC TTTTTTAAAG
551 TGATGTGACG CTTAATGAGA AGAAATCAGT AGAGAATTAC AACTGCACT
601 TCAAAAGATA CATCTAATAT CATTTTAATA ATGAAATTG AAAAAATAGT
651 GTGCTCGTTT TACAGTCTCA TTAAATGAAT TAAAATATCA GCACACATTG
701 TAGTAGGTTA TCATTGGCAG AGAAGGCTGA AATAGAAACG TTACAATGGG
751 ATGCACTGCC ATCTGAACAT TATGTCGAAG TGGAACGCGG AAACATATTT
801 CTCAGAACAA GTGGTAAAT GAAAACAGCA TCATTTGTAA AGCATTTCTT
851 TTGAGAGTGC TTCAGTTTCT TCTCCTGATG ACCTGCCATT CAGAACTGA
901 CAATGAATAA TACACTCTGA CACCAGCATT TGTCAATTG CCCAGAACCA
951 TATGAGAGTA CTCTAGACAG ATATATGTTT CGAAGTAAAC CGAATACCTG
1001 TTAAGTGTA ATCAAATCTT GTAGAAACCA TGCCATGGTT CCTTTGGACA
1051 TATACTTTGC ATGCCTGAAG CAAGTTACCT TAAGAAATCA TTCTTTTGTT
1101 TTACAAAAC TGTATTAAAA AATTAAAAAT GCAAAAAGC TTAATATTAT
1151 TAGGAATTTA TCCATAGCTT TATTTGGAAT CCAGTTTCTT TATTATGATC
1201 TATAAACATG CATCATTTGA TGGAGTTCCT TAGTGGAGAG GTGTTTTTCC
1251 ATGTTGCTAA GAAACATGCC CCAGCACCAG AAGGGATACT ACCTACCATC
1301 TTTTGGCCAT TTCTCACCGT GATTCCTACA TTGTACCTGT TTAAGTCACTG
1351 AACAGGCTT CCTTCTCTTT GTCTAGATTC TAATCAGGTG TCTTCTGGTG
1401 TGGAAGCTTT GGCTTTTATT TACACACAAC ACAGAATTAA TAAGATAGAT
1451 GCCAAGGATT TAGCAACATT TTAATTCAAC ATTATACAGG TATCAGAGTT
1501 AATGAGAATT ATGCATTAGT CTTTAAATTT GGGCAGCTTA TTCAGCTAAA
1551 ACATAGATGT CTAGCTCTTA AACACTTTGT TTTTAAATT ACTCTGAAAT
1601 TACAATAAAG TCAAAGAACT GAACTGTTTT CTTTCAAGC CAGTGCAAAT
1651 GTGCTTTAGT TATTATTTTA CTGGTGATCT AATTATGCAT TTTAATGCTT

Fig. 6

60/124

1701 TATTACTTAA TACTTATATA AGCCTAAAAT ACGTTGTTAA TGTCATAATT
1751 TCAGGGATTT TAGTATTCTT TCCATGAGTT ACCATAACTA GGTGCATATG
1801 TGTAATATATA CGTATATATC TATATCTATA TATTTATATC TATGTATATA
1851 TCAATTTATA AGACTAAATA GACTTGGCCA TATGTGTTGT TGGTTTATGC
1901 ATACATGCAC AAATATTGAG GTGTCCACAA AGTATATATG CCTGTACATA
1951 AATTACATAC TGGCTGGTGA GTGAATGTAA GCTTCTCTAA ATTGTACAAC
2001 TCTCCACAGA GTGGCACTCT AATATTGCAA AGGTACAATA TAAGCATGTG
2051 CAGAAATGAAC AGCTCTTCTA GGATCCCTAT AAAACTCCAC CCCATGTTTC
2101 TGT

Fig. 6 (cont'd)

61/124

Human genomic sequence

1 AAGCTTCATC CCAGAGGGGC ACTTGCCAGA TGCCTGCTAG AGCTCTCCTG
 51 TATGAGGAGT CTATCAACAC CTGCTGGGAG GTGTCTCCTC GTCAGGAGGC
 101 ACGGGGGTCA GGGACCCACT TGAGGAGGCT GTCTGTCCCT TAGCGGAGCT
 151 AGAACACTGT GCTCGGAGAT CCGCTGCTCT CTTCAGAGCT GGCAGGCAAG
 201 AGTGTTTTAG TCTGCTGAGC CTGCGCCAC AGCCGCCCT TCCCCAGGT
 251 GCTCTGTCCC AGGGAGATGA GAGTTTATC TGTAAAGCCC TGACTGGGGC
 301 TGCTACCTTT CTTTCAGATA TGCCCCGCC AGAGAGGAGG AATCTAGAGA
 351 GGCAGTCTGG CTACAGCAGC TTTGCCAAGC TGCAGTGGGC TCTGCCAGT
 401 CCAAAATTC CAGCGGGTTT GTTTACATTG TGAGGGGAAA AGCACCTACT
 451 CAAGCCTCAG TTATGGCAGT TGCCCCCTCC CCCACCAAGC TCCAGGGTCC
 501 CAGGTGTCCT TCAGACTGCT GTGCTGGCAA TGAGAATTC AAGCCAGTGG
 551 ATCTTAGCTT GCTGGGCTCC ACAGGGGTGG GATCCACTGA GCTAGACCAC
 601 TTAGCTCCCT GGCTTCAGCC CCCTTTCCAG GTGAGTGGAT GGTCTGTCT
 651 CACTGGCATT CCAGGTGCTA CTGGGGTATG AAAAAAAAAA CTCCTGCAGC
 701 TAGCTTGGTG TCTGCCAGT TTTGTGCTTG AAATCAGGC CCTTGGTGGT
 751 GTGGACACCC AATGGAATCT CCTGGTGTGC ATGTTGTGAA GACTGTGGGA
 801 AAAGCATAGT ATCTGGGCTG GATAGCTCCG TCCTTCAAG CACAGTCCCT
 851 CATGACTTCC CTTGGCTAGG GGAGGGAGTT CCCCACCTT TTGCACTTCC
 901 CAGGTGAGGC AACACCCAC CCGCTTCTG CTCACCTCT GTGGGCTGCA
 951 CCCACTGTCT AATCAGTCAC TGTGAGATGA GCCTGGTACC TCAGTTGGAA
 1001 ATGCAGAAAT CACCTGCCTT CTGTGTTGAT CTCACTGGGA GCAGCAGACT
 1051 GGAGCTGTTC CTATTCAGCC ATCTTTCTCA GGTCAATC ATAGATTTTT
 1101 AATTGATCCC AGCAACATGG ATTAGTAAAC AGCATATTC CAAGTGATTT
 1151 TTTTATTTT TAAGGTCAA TCTACAAAT ATTATAGTGT TATCACCCT
 1201 TAAAATTATT ACTGGTGATA CTATGTTTGT CTCTATTCAC ATTTTATTGC
 1251 TAGAAAGAAT TATAATTTGT AGATAATAAT AGTTATTTGA AATGTATTAC
 1301 ATATCCTTTT ACTTTTAAGA AGAGGTGACT TAATTATCTA GGTATACAAT
 1351 TATTTTGAGG AACTAAATG TCATGAATAG CAAATTTATC ATATTGCTTT
 1401 CCTAGGTGAA GACCCTGAAA CAAGAAGAAT GAGAACAGTT AAAACATAG
 1451 CAGACTTGAG GCAGAATTTA GAAGAGACTA TGTCCAGTCT TCGTGGGACT
 1501 CAGATAAGCC ACAGGTTTTT TTCAATTTTG CATATATTTG AGCCAATAAA
 1551 GAAAAATAA TTACAAACAA ACATTTAACT TTTCTTATAA TGACAGAGAT
 1601 GGGATTTCAG TTTCCCTTA CTATTTCTC CCTTGTTTTA TATCAAATTG
 1651 ATTGGAATT ATCCTTAAAC TGAGAATTCA CAGTATATAC CTATTTATCT
 1701 TTTATCTCTA TCTCTATCTG CTATTTATGT CTTTTTCAGT ATAATTCCA

Fig. 7

62/124

1751 GTACTGCAAC TACCACCATC ACTGTTAAGT GGATTTGTAA TACCTGTCCT
 1801 AGAAAACAGT GGCACAAGTT GCACTTGAAA TGCATCTGGG CAGGGTAGTA
 1851 GGGAGACATT CAAACATAAT TGTAAGTTAAC TTTCAGAATA GGTCTGGGAA
 1901 GGTACAGTG AGTTAAGGAT TTGTTGAAAA TGTAACAA TATGTTGTTT
 1951 TACCCAAGGT GTACTGATGG CCTTTCTTTT GAAACAAAC GAAAAGCTAT
 2001 AAAATGTATG CCCCTTTCCA CAATTTGACC TCAAAATGAA TATAGAGTTT
 2051 AGCTTTCGGG AAGATGACGT GTTTATAAGA GATGACCCCTC AACTCCAGCC
 2101 TTTTCTGTCT TCATGCATTC TAGATTATGG CCCTAAGTGA ACCAGAGTAT
 2151 AGTTATTTCT CCATTTTAT TACAGCACC CTGGAGACAA CATTTGACAG
 2201 CACTGTGACA ACAGAAGTTA ATGGAAGGAC CATACCCAAC TTGACAAGTC
 2251 GACCCACCCC CATGACCTGG AGGTTGGGCC AGGCATGTCC GCGACTTCAG
 2301 GCGGGAGATG CTCCCTCCCT GGGTGCTGGC TATCCTCGCA GTGGTACCAG
 2351 TCGATTCATC CACACAGACC CCTCGAGGTT CATGTATACC ACGCCTCTCC
 2401 GTCGAGCTGC TGTCTCTAGG CTGGGAAACA TGTCACAGAT TGACATGAGT
 2451 GAGAAAGCAA GCAGTGACCT GGACATGTCT TCTGAGGTCG ATGTGGGTGG
 2501 ATATATGAGT GATGGTGATA TCCTTGGGAA AAGTCTCAGG ACTGATGACA
 2551 TCAACAGTGG GTAAGTAACC CTGTTCTCCG TCAGCATTGT GTGAAGAGGG
 2601 GAGGTGGTCT ACTATAATGC ATTCATAATA AACAAATGTG TAAGTTTGCC
 2651 CAGAAAGTCA TGAGAACATA TGAGATATCT GAGGTTATTC AGAGTGTGA
 2701 AGGGCCCTTC CTCTGCTCAT TCATGGAGAG TAAAGAATCC AAGATTTCTA
 2751 TAAATTCATT ATAAGCCGCT AAGTTTTTCT GTTGTGAGA GAAACACATG
 2801 TGGCTTCTGT TTTTCAGAGT GATTTTCACA TGCTTCTTAA GTAACAGATT
 2851 TTGTAGTTAA GGACGTGGGA AGGAGACAGG AGGAGTTTGG CTGATTTGCT
 2901 TGATTTTTTT TTTCTTTTTT AGCTTGTTAG AAGCGGCCCTG TAAGTCTTTT
 2951 GAGAAACAAA TATTTTCTTA CTGTCTTCAA TTATGCATCC CCAATTTAAC
 3001 TTGAGGGAAA AATCACTTTG GAGTTGAAAG TTCACTCTA TTCATTTTCT
 3051 TTTGATGGTA TCAGATTTCA ATACATCTCA GACCCTGTTT TTCTTCTGTG
 3101 TCCTATTACA TTCCAAAACA TGTTGTGATT GTAAACTCT TAGAGTATAT
 3151 TAACAATTTG GGATATTTGG CATAATCAGA GAATAGGTCC AAAAGGAGGC
 3201 AATAGGATAT TCTATTAATA ATTGTAATTG CCATTTTATG CATTTCTGT
 3251 TATGTACTAT GCTCTTGTC AAGTCTTTGA AGATAGTGT TTACTTTTCC
 3301 TTCCCACCAC CAGCAATGTT TATGAGGTAG ATGTTTTTAT ACATGTTCTA
 3351 TGGATAAGGA AACTGAGTCT AATTGGCCCC GGCTGGGAAC TAACGCTAGG
 3401 GAAACGGCAG ACCTGCATTA GAACTCAGCT ATGTCTGACT TCAAACACAG
 3451 GCTCAGTAAT ATGTGGAAAA GCTTCCCAAT TAACTTTGTC TATAAACTTT

Fig. 7 (cont'd 1)

63/124

3501 GTGTGAGTCT GGATTTTGAC TTACTCTTTG TCTTTACGCA TCTGAGAGGA
3551 CCCATGTAGG AAATAATTCT TCTATATAAG TGACCCCTTC TGACTTCATT
3601 CATGAAAAGC TTATGTTTGA AGGGTGACAC GACCTAAAAA AGAGTACAAA
3651 ATAGCTTTTG ATTACATTTA TAGCTTTGCT CTGATATCCT AATACCTACT
3701 AGTCCATTCC TGGTATCCAC CCTACCTGAC TTTCTAAAAA TTTAGAATTA
3751 TAGAGACTAA TTATGATTAA TTAAGATAGG TTGTTGTTCA GTTGCCACTG
3801 GATTCAGAGT GCCTAGTTTG AATCTCTCCC ATTCATATC TGTGGACCCC
3851 TTCGGAACCT AACGTATCCA AATTAGTTTT TGTCATCTAG AATAAGGATA
3901 AAATTGTACC ATCTTCATGA AGTTGTTAGG ATCATCCACA AATTTTAGTT
3951 TGCGCAATGC TTGGCATGAT ACAAGCACTC AATAAATTTA TCATCTTCCT
4001 CTTTATCATC ACTATTACAT TTATTATCAT TAATAACCAT ACCAATTTTT
4051 GGTGTTGTT AGTTATAATT ATCATTTTTG TATGTATTTA ACATAGCCTA
4101 GGAGGCAATG CCCAGTTCAG AAAACATAAT GGCAAAGCAA GAGTGTCTAA
4151 GGCACACTCT TTCTCCCATC TCTCTCTTCT TTCTTCTCCA TTCTTTCCAC
4201 TCTATCCCT CTTCTCTTTT TTTTCTCAAT CTCCTTAGAT GTGGACATAT
4251 GTGTGAATTC

Fig. 7 (cont'd 2)

64/124

Human genomic sequence

1 TGTGGGTGTG GGTGTGAAGC ATGTGTATGT GTGTGTGTGA AGCATCTCCC
51 CACCTGTAAT GTAAGTCCAT GAGTGCAGAA TTTTGTACAT ATTCTTTACG
101 TGTTGAGTTT TAACAAATGT TTGTGGAGTG AATGAACAAA TTAATGAATA
151 TAGGCTATTT ATTAATTAGG CAATATAGTC ACATAGGCTG GCAATCGCAT
201 CTAATTAAAT AGAGTGGTAA ATGAGTTCCA GAAAGAACTA AGGTACTACA
251 AGGATGTTAT GAAAGAGAAA AATGAGTTAT GTGAAAAATA GGAGACAGTG
301 ATAAGAGGGA AAGAATCCCA AAGTGTGGGC CACATTTTGA AACTAATGAC
351 CTATTATCTT ATTATTGTTA GCTGAAAGTA GAAAACGTCA TGGGAGGGAA
401 TATCTGCTAG TTTTGTGTAA AGGATGTTGT GATGGCAGAA CCAAGAAATG
451 AACACAAGGT GACTTTGGTT TGGGGACAGT GGGATAATCA ACTCTCCTTG
501 CTCCATCAGG GCCCCAGACT GGGCTCTGGC AGAGGAACTC AGAACAACGT
551 AAAGACCTAG ATAGGTATCT AATAAATTGG GACCTGTGAA AACAGTGCCT
601 CTTAAAGTGT GGTACCTGGA CCAGCAGCAG CAGCAGCAGC AGCCATTGAA
651 ACTTCATAGA AAGACAGATT CTCAGCTTCA TCCAAGACTT ACTGAATTAG
701 AATATCTCAA GGTAAGGCCT GGTAATCTGA GCTTTAACTA GCCCTCAAGG
751 TGATTCTTAA GTTCAAGCAT CACTATATTA AGTTGAACAA ATAGATGCCA
801 GGCCTATAAA TACATGTAAC GCCTAGCATA AATATTTCAA CATTAATAAT
851 GACATTTTCA AGTTCTTATT TACCCTATTA GCTGTGTTCT GTCAAGATAA
901 TGAGAATATT GATATGTTAG AATACACTGA TGCACTAATT TTTAAATTAG
951 ATCAAATAAT GACTTGTTAT ACCTGAAATA AATTGGTTCA GCTTGGTAGA
1001 TGCAGTTTTT GAGAATTATA TAAGTCATTT TAAAAGAAT AATTTTAACT
1051 TGAGCTGCTT GCATAAATTA AATTGCAAAA AGGTCATAGT ATAAATCCTC
1101 CTATTAGCAG AGATAGAAGG TTTTAAAAA AATTACAGAT AAGTCTGAAG
1151 GTCTTTTAAA ATCTTATATT CAGGAAGTGA CTCGGGATGT ATATCATTTT
1201 AAAATACATG GTCTTAAATG TTGTAGTTGT ATGACTCTTT CAGTTAATTT
1251 AAAATACTTC CTTCTATGAA AAATTGTTTC AAAAATTTTT CTAAATTCTG
1301 TTATCCATTT CAAGTAGGAT AGGCAAGAAC AGATATAAGA TACTACTTTT
1351 TTGTTTCATG TTAATAAAAA AAAAATTACT GTAATTGAGA TCATGTAAAA
1401 ACATGTTTCC TGTCTATTTG TCTTAACCTT TTAATCCTGG CACCTTAAAT
1451 TTGACATAGT AGGAATTAGA AGACAATTGC AGAAAATGTC AACTGGGGAA
1501 ATTTTATTCT ACTAAAACT ATGTCCATAC AACATAGCAA ATCACATTTT
1551 AAAGGCCAAA AAGTCTTTCA TAGCAATTTT TCAGATTATT TTCAAAGCAT
1601 ATCTTCTCTC TGCTCCTGCA GCATGCCGTT GATTTTTCTG TTATGCAGTC
1651 ACATAAGTAA TTACATGTTT ACATGTCTAT TTCCTCATA GAACACGAAA
1701 CAGTTAAATG TAGAATAATA TCCAATCCAT CTTTTTATCA CCAGTAGCTA

Fig. 8

65/124

1751 GCATACTGTA GGAACCTCAAT AAATATATCA GATAAATTGT GGAAATAACC
1801 ATATCAGCTT ATAACATATA GAAATGTGAG TTTAAAAAGA AAACAATTAT
1851 ACATATGAAA AAATTTTTAT ACCATTTTTT TAAAGACCTT TCAGATGTCA
1901 TACAGTTTGG ACTTTTCCAG TGTTCCTTGT ATCATGAGAC AATAGTAGAC
1951 ATTGTAAATC AAAAATAGTT TTCTGGGGTT GTGTACATTT GAAAAAACTG
2001 AATATCATAT CTGTTCTTAG AGAGTAATGA TGGATATTAA CATATCAAAG
2051 GTACAGAGAA GTCTTAAAGT TCAAAGTAAC ATCTGCTTAA TTGTATTTAA
2101 TTCAGTGCTC CATGAGCTTT TTTATCACTG ATTCCCTCCC TTTTTTCTCT
2151 TATGATAATA ATTAACCTGT TCCTGTAGCA TTTTAAGAAA TGTTGATTTA
2201 GTTGAATGCC TTCACTTCTC CAATATAATA GCAGAACTC AGAAATATTT
2251 ATTTACCCAG AATCATGCAG CTAATAGTAC AAGGATTCAG GTCTTTTACT
2301 TCCTATTTTG TGGTTCCCAA CTACTTTTGC CAAAGGTCTT TTAAATAATA
2351 TGAAACATAT TAGTGATTGA TTCATTATAG TAAATGGGTA AATGATAAGG
2401 CTTGCAATAA TTCACTGACA AGAAAGCTT

Fig. 8 (cont'd)

66/124

Murine cDNA sequence

```

1   AAGCCACAGCACCCCTGGAGACAACCTTTGATACGACTGTGACAACCTGAAGTGAATGGAAG
   S H S T L E T T F D T T V T T E V N G R

61  GGCCATCCCCAACCTGACAAGCCGACCTTCCCCCATGACCTGGAGACTGGGTCAAGCGTG
   A I P N L T S R P S P M T W R L G Q A C

121 CCCTCGTCTACAGGCTGGAGATGCCCCCTCCATGGGCGCTGGATATTCTCGAAGCGGTAC
   P R L Q A G D A P S M G A G Y S R S G T

181 CAGCCGATTTCATCCACACGGATCCCTCCAGGTTTATGTATACCACGCCCTCTCCGCCGAGC
   S R F I H T D P S R F M Y T T P L R R A

241 TGCTGTCTCGCGTCTGGGAAACATGTACAAATAGATATGAGCGAGAAAGCAAGCAGTGA
   A V S R L G N M S Q I D M S E K A S S D

301 CCTGGATGTGTCTTCTGAAGTGGATGTTGGTGGATACATGAGCGATGGTGATATCCTTGG
   L D V S S E V D V G G Y M S D G D I L G

361 GAAGAGTCTGAGAGCGGATGATATCAACAGTGGGTACATGACAGATGGTGGGCTCAACCT
   K S L R A D D I N S G Y M T D G G L N L

421 ATATACCAGAAGTCTTAACCGAGTCCCGGACACAGCAACTTCCAGAGATGTCATACAGAG
   Y T R S L N R V P D T A T S R D V I Q R

481 AGGCGTTACGATGTGACAGTGGACGCAGACAGCTGGGATGACAGCAGTTCTGTGAGCAG
   G V H D V T V D A D S W D D S S S V S S

541 TGGCCTCAGTGACACACTTGATAACATTAGCACAGATGACCTCAACACCACGTCTCCAT
   G L S D T L D N I S T D D L N T T S S I

601 CAGTTCTTACTCCAACATCACTGTCCCCTCCAGGAAGAACAACCTCAGCTGAAAACAGATGC
   S S Y S N I T V P S R K N T Q L K T D A

661 GGAGAAACGTTTCGACAACAGATGAGACCTGGGATAGTCCTGAGGAGCTGAAGAAAGCCGA
   E K R S T T D E T W D S P E E L K K A E

721 GGGAGATTGTGACAGCCATGGTGACGGAGCCGCCAAGTGAAGGGTGCTACTTCTGGACT
   G D C D S H G D G A A K W K G A T S G L

781 TGCTGAAGACTCGGAGAAGACAGGGCAGAAAGCCAGCCTGTCTGTGTCTCAGACAGGCTC
   A E D S E K T G Q K A S L S V S Q T G S

841 CTGGAGGAGAGGCATGTCTGCCCAGGGAGGAACCTCCAGCTACAGCTAGGCAGAAAACCAG
   W R R G M S A Q G G T P A T A R Q K T S

901 CACAAGTGCACTCAAGACCCCTGGGAAGACAGATGATGCCAAAGCTTCCGAGAAAGGGAA
   T S A L K T P G K T D D A K A S E K G K

961 AACTCCTCTCAAAGGATCATCCTTGCAAAGGTCTCCTTCAGATGCAGGGAAAAGCAGCGG
   T P L K G S S L Q R S P S D A G K S S G

1021 GGATGAAGGGAAAAAGCCACCGTCAGGCATTGGAAGATCGACAGCCAGCAGTTCTTTTGG
   D E G K K P P S G I G R S T A S S S F G

1081 ATACAAGAAGCCAAGTGGTGTAGGGGCTTCCACTATGATTACCAGCAGCGGTGCCACCAT
   Y K K P S G V G A S T M I T S S G A T I

1141 CACAAGCGGTTTCAGCTACACTGGGGAAAATCCCCAAATCCGCTGCCATTGGTGGGAAGTC
   T S G S A T L G K I P K S A A I G G K S

1201 CAATGCAGGAAGGAAAACCAGCCTGGACGGGTCCAGAATCAAGATGATGTTGTCTCTGCA
   N A G R K T S L D G S Q N Q D D V V L H

1261 CGTGAGCTCGAAGACCACCCTCCAGTACCGTAGTTTGCCCCGCCCTTCTAAGTCCAGCAC
   V S S K T T L Q Y R S L P R P S K S S T

1321 CAGCGGAATCCCTGGGAGAGGTGGCCACAGGTCGAGCACCAGCAGCATTGATTCCAATGT
   S G I P G R G G H R S S T S S I D S N V

```

Fig. 9.

67/ 124

1381 CAGCAGCAAGTCAGCTGGGGCCACCACCTCCAACTGAGAGAACCGACTAAGATCGGCTC
S S K S A G A T T S K L R E P T K I G S

1441 AGGGCGCTCGAGTCCAGTCACTGTCAACCAACAGACAAAGAGAAGGAGAAAGTAGCAGT
G R S S P V T V N Q T D K E K E K V A V

1501 GTCAGATTCAGAGAGCGTTTCCTTGTCAGGTTCCCCCAAATCCAGCCCCACCTCTGCCAG
S D S E S V S L S G S P K S S P T S A S

1561 TGCCTGTGGGACTCAAGGGCTCAGACAGCCAGGGTCCAAATATCCAGATATTGCCTCGCC
A C G T Q G L R Q P G S K Y P D I A S P

1621 CACATTTGCAAGGTTGTTTCGGTGCCAAGGCAGGCGGCAATCTGCCTCCGCACCTAATAC
T F R R L F G A K A G G K S A S A P N T

1681 TGAGGGGGCGAAGTCCCTCCTCAGTAGTGCTCAGCCCTAGTACCTCTTTAGCCCGACAAGG
E G A K S S S V V L S P S T S L A R Q G

1741 CAGTCTGGAGTCACCGTCGTCGGGTACGGGAAGCATGGGCAGTGCTGGTGGGCTGAGTGG
S L E S P S S G T G S M G S A G G L S G

1801 CAGCAGCAGCCCTCTCTTCAATAAACCCCTCAGACCTAACTACAGATGTTATAAGCTTAAG
S S S P L F N K P S D L T T D V I S L S

1861 TCACTCCTTGGCTTCCAGCCCAGCGTCGGTTCCTCTTTTACATCCGGTGGGCTTGTGTG
H S L A S S P A S V H S F T S G G L V W

1921 GGCTGCCAATCTGAGCAGTTTCTCTGCCGGCAGCAAGGACACTCCAAGTTACCAGTCCAT
A A N L S S S S A G S K D T P S Y Q S M

1981 GACTAGTCTCCATACGAGCTCTGAGTCCATTGACCTGCCCCCTCAGCCATCATGGCTCCCT
T S L H T S S E S I D L P L S H H G S L

2041 GTCTGGACTGACCACAGGCACTCACGAGGTGCAGAGCCTGCTCATGAGAACGGGTAGTGT
S G L T T G T H E V Q S L L M R T G S V

2101 GAGATCTACTCTCTCAGAAAGATACACCCCATCATCTCGGCAGGCCAACCAAGAAGAAGG
R S T L S E R Y T P S S R Q A N Q E E G

2161 CAAAGAGTGGCTGCGATCGCATTCCACTGGCGGGCTGCAGGATACTGGCAACCAGTCTCC
K E W L R S H S T G G L Q D T G N Q S P

2221 CTTGGTCTCCCTTCTGCCATGTCATCGTCAGCCACCGGAAAATATCACTTTTCCAACCT
L V S P S A M S S S A T G K Y H F S N L

2281 GGTGAGTCCCACCAACCTCTCCAGTTTAACTGCCTGCACCCAGTATGATGCGCTCCAG
V S P T N L S Q F N L P A P S M M R S S

2341 CAGTATCCCCGCCAGGACTCCTCCTTCGACCTCTATGATGATGCCAGCTTTGCGGTAG
S I P A Q D S S F D L Y D D A Q L C G S

2401 TGCAACTTCCCTGGAGGAAAGGCCACGGGCCGTTAGCCACTCCGGCTCATTCAGAGACAG
A T S L E E R P R A V S H S G S F R D S

2461 CATGGAGGAAGTTCATGGCTCTTCACTGTCATTGGTCTCCAGCACATCATCCCTTTACTC
M E E V H G S S L S L V S S T S S L Y S

2521 TACGGCTGAAGAGAAGGCTCATTCAGAGCAATCCATAAGCTACGGAGAGAAGTGGTTGC
T A E E K A H S E Q I H K L R R E L V A

2581 CTCCCAGGAGAAAGTCGCTACCCTCACGTCTCAGCTGTCAGCAATGCTCACCTTGTTAGC
S Q E K V A T L T S Q L S A N A H L V A

2641 AGCTTTTGAAGAGAGTTTAGGGAATATGACTGGCCGTTTGCAAAGTCTAACCATGACAGC
A F E K S L G N M T G R L Q S L T M T A

2701 GGAACAAAAGGAATCTGAGCTTATCGAACTGCGGGAACCATGAAATGTTGAAGGCCCA
E Q K E S E L I E L R E T I E M L K A Q

Fig. 9 (cont'd 1)

68/124

2761 GAACTCTGCTGCCCCAAGCAGCCATTCAGGGAGCACTGAATGGCCCAGACCACCCTCCCAA
N S A A Q A A I Q G A L N G P D H P P K

2821 AGATCTCCGCATCAGAAGACAGCACTCCTCTGAAAGTGTTCCTAGTATCAACAGCGCAAC
D L R I R R Q H S S E S V S S I N S A T

2881 GAGCCATTCCAGCATTGGCAGTGGTAATGATGCTGACTCCAAGAAA
S H S S I G S G N D A D S K K

Fig. 9 (cont'd 2)

69/124

Murine genomic sequence

1 GGGATGAAGG GAAAAAGCCA CCGTCAGGCA TTGGAAGATC GACAGCCAGC
51 AGTTCTTTTG GATACAAGAA GCCAAGTGGT GTAGGGGCTT CCACTATGAT
101 TACCAGCAGC GGTGCCACCA TCACAAGCGG TTCAGCTACA CTGGGGAAAA
151 TCCCCAAATC CGCTGCCATT GGTGGGAAGT CCAATGCAGG AAGGAAAACC
201 AGCCTGGACG GGTCCCAGAA TCAAGATGAT GTTGTCTGTC ACGTGAGCTC
251 GAAGACCACC CTCCAGTACC GTAGTTTGCC CCGCCCTTCT AAGTCCAGCA
301 CCAGCGGAAT CCCTGGGAGA GGTGGCCACA GGTCGAGCAC CAGCAGCATT
351 GATTCCAATG TCAGCAGCAA GTCAGCTGGG GCCACCACCT CCAAACCTGAG
401 AGAACCGACT AAGATCGGCT CAGGGCGCTC GAGTCCAGTC ACTGTCAACC
451 AAACAGACAA AGAGAAGGAG AAAGTAGCAG TGTCAGATTC AGAGAGCGTT
501 TCCTTGTCAG GTTCCCCCAA ATCCAGCCCC ACCTCTGCCA GTGCCTGTGG
551 GACTCAAGGG CTCAGACAGC CAGGGTCCAA ATATCCAGAT ATTGCCTCGC
601 CCACATTTCTG AAGGTAAGGG TATGTAAAGA GATGTTGGGA AAACATAAAA
651 GGTAGTATAT AGCATGTATT TATTCTGTAC GAAACTATTT TCATGTATTCT
701 TAAATATTCT AAGATTCTGT ATCTTATACT TGTCTAAAAT ATAGTGATTT
751 TATTTTGCTG ATTGCACCTG TTGCTAGTGT AAAAGCATTG CTCATTTAGA
801 GAGTGGTTAG CCTTTCAGCT ATACAGCCAG TGTGACACTA AAATACAGAT
851 ACCACTTGTA GCGGGCATAA AACCACATGA CTGACTATTC ATAGAAATAA
901 AGTGATAGCT TGTAAGATA TTAGTGATT TCCACCTCTC CTTTCCAGAA
951 TTAAAAAAG CAAATTGCAT AGATCTTTAT AAACACATTT ACTTCTAGTG
1001 TATGTTATCT TGTTGACTCT TAATGAAATG GCAGTTATGA ATATAGATGA
1051 TATATTCTTT CTAACAGTTT ATAAGAGACC AATTTATACA GTACCAGATC
1101 TTAACATAGT AACAATAACA GCAACAAAAA CAACCCAAAA AGCTATCAAA
1151 GTATGGTCTG ATTGCAGAAT TTGAAAACAT TTACATGTTT GACATAGGAC
1201 AAGAACTCAG GAGTGAGGTG ACTTTTATA AGTCTTCATC AATGTCCTTT
1251 TACAGGAACC AGGAAGCATA TCTGATATAT GTGTCAGGAT TATCACTTTA
1301 TTAATTATGT GAAATTCTGT TTAGAAATCT ACCTGATTTT AAATACTTTA
1351 ATATAGTAGG GGTCAAAATT AGTTAATGAG TTAAGACAAG TTGTAAATA
1401 ATCCTGGCTC TGTTTTCTCA TCTTCAAAAT GATAGAGTAT AATTTATCAC
1451 CTCTTGTTAA ATATTTTCAGG TTTGTGTTTA TTCTCTGAT AACTTTGATC
1501 TCTTAGAAGA GTCCTGAAGA ATTTACATTA AGTAATCTTA GAAACATAAC
1551 TATTTGAGAA ACAGTAGTCA AATTTTGTC TTAGAAGTAT TAACTCTGAA
1601 GAATGATTTG AAGTGACAGT TCTTAGAAAG AATAAATTAT AGCTTGTAGC
1651 AAGAGTAAAT ATTTTCACTG CTTGTGTGAG AGCCAAGAGC GCCCTCTTGT
1701 GGCCCATTTAC CTATGAAACA ATTTCTCATA TTCGCCCTAG AAATCTTCCA

Fig. 10

70/124

1751 CTGCAGGAAA TAATGGATTT CATTGCCTCT GAATTAGTAA CCATTCTGCC
1801 ATTTCTTCAT ACCATTTTAT TTCCATACTT GCATAAATTT GATTATGTCA
1851 TCTGCTTCAT TTACAAAACCT AAAATGTTTT CTGAGCTAAA CTCCAGTAGC
1901 TAACTTAGTA CAAATGGTAT TTTTAAATCA CTGCTATAAG TATATATATT
1951 TGAATAGCTC TGGCAACGGA CGGAAATCCC TATGGTCTTT CCATGGGAAG
2001 ATACAAACCA ATCCATAAGT TGTCCAGCAA TATCCAATAT TTCCAGCCCA
2051 GCCAGTCAGG CCTCTTAAAC ATTACCTTAC ATATTTGAAC CTTTCCTTAA
2101 ATGTCCCCTT TAGACAATCT ATTTTTTAAA AAGATGAAAA TCCATTTAAG
2151 CATCATATAT CGAATGCGTA GAAGTTGTTT CATTATAATG GTTCTGCAGA
2201 TAGGTAATGC CAAAACGGCC AAAATATTTG ATCACTAGAA GCGTAAAAGT
2251 CAAGTACAAT CATGTTGACT TTTTTTCCAA GGTGGGTTCA CTGCTGCCCA
2301 CCTTG GTTCC AGGCCAGTGC TTA CTTAAGA TATCGTAAGT GATTTTTTTT
2351 TAATTTTTAA TTTTTTAGTA GTTGGTTAAT CAAAAGCCAG TCATGTCACC
2401 TTCAGGAACA TAGAGGCTGG ACGTGCTTGG CAGCTCACGA CTCCAAAGCA
2451 CACTTGGCTC TGTGGACTGA AACCC TAGGA AACGTGGATG TGAGTCTCTT
2501 GGAACAACCTC AAGTTGTTAT TTGTTTTTCT TTTAGGTTGT TCGGTGCCAA
2551 GGCAGGCGGC AAATCTGCCT CCGCACCTAA TAC

Fig. 10 (cont'd)

71/124

T2HC

Homologous human cDNA

1 GGATCAGCTTCGGGAGACCATGCACAACATGCAGTTGGAGGTGGACCTGCTGAAAGCAGA
D Q L R E T M H N M Q L E V D L L K A E

61 GAATGACCGACTGAAGGTAGCCCCAGGCCCTCATCAGGCTCCACTCCAGGGCAGGTCCC
N D R L K V A P G P S S G S T P G Q V P

121 TGGATCATCTGCATTATCTTCCCCACGCCGCTCCCTAGGCCTGGCACTCACCCATTCCTT
G S S A L S S P R R S L G L A L T H S F

181 CGGCCCCAGTCTTGACAGACACAGACCTGTCACCCATGGATGGCATCAGTACTTGTGGTCC
G P S L A D T D L S P M D G I S T C G P

241 AAAGGAGGAAGTGACCCTCCGGGTGGTGGTGAGGATGCCCCCGCAGCACATCATCAAAGG
K E E V T L R V V V R M P P Q H I I K G

301 GGACTTGAAGCAGCAGGAATTCTTCTGGGCTGTAGCAAGGTCACTGGAAAAGTTGACTG
D L K Q Q E F F L G C S K V S G K V D W

361 GAAGATGCTGGATGAAGCTGTTTTCCAAGTGTTC AAGGACTATATTTCTAAATGGACCC
K M L D E A V F Q V F K D Y I S K M D P

421 AGCCTCTACCCTGGGACTAAGCACTGAGTCCATCCATGGCTACAGCATCAGCCACGTGAA
A S T L G L S T E S I H G Y S I S H V K

481 ACGAGTGTGGATGCAGAGCCCCCGAGATGCCTCCTTGCCGTCGAGGTGTCAATAACAT
R V L D A E P P E M P P C R R G V N N I

541 ATCAGTCTCCCTCAAAGGTCTGAAGGAGAAATGCGTCGACAGCCTGGTGTTCGAGACGCT
S V S L K G L K E K C V D S L V F E T L

601 GATCCCCAAGCCGATGATGCAGCACTACATAAGCCTCCTGCTGAAGCACC GGCGCCTCGT
I P K P M M Q H Y I S L L L K H R R L V

661 CCTCTCGGGCCCCAGCGGCACGGGCAAGACCTACCTGACCAATCGCTTGGCCGAGTACCT
L S G P S G T G K T Y L T N R L A E Y L

721 GGTGGAGCGCTCTGGCCGTGAGGTACAGAGGGCATCGTCAGCACCTTCAACATGCACCA
V E R S G R E V T E G I V S T F N M H Q

781 GCAGTCTTGCAAGGATCTGCAACTGTATCTTTCCAACCTAGCCAACCAGATAGACCGGGA
Q S C K D L Q L Y L S N L A N Q I D R E

841 AACAGGAATTGGGGATGTGCCCTGGTGATTCTATTGGATGACCTGAGTGAAGCAGGCTC
T G I G D V P L V I L L D D L S E A G S

901 CATCAGTGAGTTGGTCAATGGGGCCCTCACCTGCAAGTATCATAAATGTCCCTATATTAT
I S E L V N G A L T C K Y H K C P Y I I

961 AGGTACCACCAATCAGCCTGTAAAAATGACACCCAACCATGGCTTGCACTTGAGCTTCAG
G T T N Q P V K M T P N H G L H L S F R

1021 GATGTTGACCTTCTCCAACAACGTGGAGCCAGCCAATGGCTTCCTGGTTCGTTACCTGAG
M L T F S N N V E P A N G F L V R Y L R

1081 GAGGAAGCTGGTAGAGTCAGACAGCGACATCAATGCCAACAAGGAAGAGCTGCTTCGGGT
R K L V E S D S D I N A N K E E L L R V

1141 GCTCGACTGGGTACCCAAGCTGTGGTATCATCTCCACACCTTCCTTGAGAAGCACAGCAC
L D W V P K L W Y H L H T F L E K H S T

Fig. 11

72/124

1201 CTCAGACTTCCTCATCGGCCCTTGCTTCTTTCTGTCGTGTCCCATTGGCATTGAGGACTT
S D F L I G P C F F L S C P I G I E D F

1261 CCGGACCTGGTTTCATTGACCTGTGGAACAACCTCTATCATTCCTTATCTACAGGAAGGAGC
R T W F I D L W N N S I I P Y L Q E G A

1321 CAAGGATGGGATAAAGGTCCATGGACAGAAAGCTGCTTGGGAGGACCCAGTGGAAATGGGT
K D G I K V H G Q K A A W E D P V E W V

1381 CCGGGACACACTTCCCTGGCCATCAGCCCAACAAGACCAATCAAAGCTGTACCACCTGCC
R D T L P W P S A Q Q D Q S K L Y H L P

1441 CCCACCCACCGTGGGCCCTCACAGCATTGCCTCACCTCCCGAGGATAGGACAGTCAAAGA
P P T V G P H S I A S P P E D R T V K D

1501 CAGCACCCCAAGTTCTCTGGAATCAGATCCTCTGATGGCCATGCTGCTGAAACTTCAAGA
S T P S S L D S D P L M A M L L K L Q E

1561 AGCTGCCAACTACATTGAGTCTCCAGATCGAGAAAACCATCCTGGACCCCAACCTTCAGGC
A A N Y I E S P D R E T I L D P N L Q A

1621 AACACTTTAAGGGTTCGGCAATCACTGTCACCCCGGACAGCAGAACGCTGGCATCAGCT
T L *

1681 ATCTTAGCTCCTCCTCTCCCCTCTCCTCTTTCAGAGCACTGGCTCTCCAGCCCCAGGAGG

1741 AGAACAGGAGGGGAGGAGGAGATGAAAGAGGAGGGACAGGTTCTTGGTGCTGTACCTTTGA

1801 GAACTTCCTAGGAAGGAATGGTGGGGTGGCGTTTGGGAACTTGTGCCCCCTAAACACATT

1861 TACTGGCCTCCTCTAATGACTTTGGGGAAAAGATGATTCTGGGTCTTTCCCTTGACTTCT

1921 TGTTTCAATTACAACTCCTGGGCTTTCTGGGGAGGGGTTTCAGAAAACATCAAAACACTG

1981 CAGCAGTTCCTAAATGATTCTCACAAGCAACCCTGAGAGAGACAGTCTTGTGAGGGAGAT

2041 CTGGGGGAGGCAGGAAGCTCCTCAGATTTTCTCACAGACCCCTTCCCAATTCCATCACCAC

2101 TGCCAACAACCTCCTCCCCAGAGATCTGGCTGGAGCCCAGAAAAGAAGCATGTGGTTTAA

2161 AAAAAATGTTTAAATCAATCTGTAAAAGGTAAAAATGAAAAACAAAAACAAGCAAACAAAC

2221 AAAAAACAATGGAAAAGATGAAGCTGGAGAGAGAGGAACCAAGTTGCCAAGGTAGAGAGCT

2281 GCCCGCTCCTGCCCTCTGGATGACATAGGGGACATCAACAAGACGGCTGCCAACCTGAGA

2341 AGTCACCAAACCACAAAAATAACCTTACAGCCTTCAGGGAAAGACTACCAGCTCTGTCTT

2401 TCTACCCTCTAATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAA

Fig. 11 (cont'd)

73/ 124

Homologous murine cDNA sequence

1 GAACTATGGGAAAAAGAGATGAAGCTCACGGATATCCGGTTGGAGGCCCTCAACTCTGCC
E L W E K E M K L T D I R L E A L N S A

61 CACCAGCTGGACCAGCTTCGGGAGACCATGCACAATATGCAGTTGGAGGTGGACCTGCTG
H Q L D Q L R E T M H N M Q L E V D L L

121 AAAGCAGAGAATGACCGGCTGAAGGTTGCCCCCGCCCCCTCCTCAGGCTGCACTCCAGGG
K A E N D R L K V A P G P S S G C T P G

181 CAGGTCCCTGGGTCATCGGCTCTGTCGTCGCCCTCGACGTTCCCTGGGCCTTGCACTCAGC
Q V P G S S A L S S P R R S L G L A L S

241 CATCCTTTTCAGTCCTAGTCTCAGACACAGACCTCTCACCCATGGATGGCATCAGCACC
H P F S P S L T D T D L S P M D G I S T

301 TGTGGTTCAAAGGAAGAGGTGACCCCTGCGGGTGGTGGTCCGGATGCCGCCCCAGCACATC
C G S K E E V T L R V V V R M P P Q H I

361 ATCAAAGGGGACTTAAAGCAGCAGGAGTTCTTCTGGGTTGCAGCAAGGTCACTGGCAAA
I K G D L K Q Q E F F L G C S K V S G K

421 GTTGACTGGAAGATGCTGGATGAAGCCGTTTCCAAGTGTTCAGGACTACATTTCTAAA
V D W K M L D E A V F Q V F K D Y I S K

481 ATGGACCCAGCCTCAACCCTGGGACTGAGCACTGAGTCCATACATGGCTATAGCCTCAGC
M D P A S T L G L S T E S I H G Y S L S

541 CACGTGAAACGAGTGTGGATGCTGAGCCCCCAGAGATGCCTCCTTGCCGCCGAGGTGTC
H V K R V L D A E P P E M P P C R R G V

601 AATAACATATCAGTCGCTCTCAAAGGTCTGAAAGAGAAGTGTGTCGACAGCCTGGTGTTC
N N I S V A L K G L K E K C V D S L V F

661 GAGACGCTTATCCCCAAGCCCATGATGCAGCACTACATCAGCCTCCTGCTCAAGCACCGG
E T L I P K P M M Q H Y I S L L L K H R

721 CGCCTGGTGTCTCTCCGGCCCCAGTGGCACCGGCAAGACCTACTTGACCAATCGGCTAGCC
R L V L S G P S G T G K T Y L T N R L A

781 GAGTACCTGGTGGAGCGCTCCGGCCGCGAGGTACGGATGGCATCGTCAGCACTTTCAAC
E Y L V E R S G R E V T D G I V S T F N

841 ATGCACCAGCAGTCTTGCAAGGATCTGCAACTGTACCTCTCCAACCTAGCCAACCAGATA
M H Q Q S C K D L Q L Y L S N L A N Q I

901 GACCGGGAACAGGGATAGGGGATGTGCCCTTGGTGTCTCCTGATGATCTGAGTGAA
D R E T G I G D V P L V I L L D D L S E

961 GCAGGCTCCATCAGTGAGCTGGTCAATGGGGCCCTCACCTGCAAGTATCACAAATGTCCC
A G S I S E L V N G A L T C K Y H K C P

1021 TACATTATAGGTACCACCAATCAGCCTGTAAAAATGACACCAACCATGGCTTGCACTTG
Y I I G T T N Q P V K M T P N H G L H L

1081 AGCTTCAGGATGCTGACCTTCTCGAACAATGTGGAACCAAGCAATGGCTTTCTGGTCCGT
S F R M L T F S N N V E P A N G F L V R

1141 TACCTGCGGAGGAAGTTGGTAGAGTCAGACAGTGACGTCAATGCTAACAAGGAAGAGCTG
Y L R R K L V E S D S D V N A N K E E L

1201 CTTGCGGTGCTGGACTGGGTGCCCCAAGCTGTGGTATCACCTCCACACCTTCCTGGAGAAG
L R V L D W V P K L W Y H L H T F L E K

1261 CACAGCACCTCGGACTTCCTCATTGGCCCTTGCTTCTCCTGTCTGTCCCATTTGGCATC
H S T S D F L I G P C F F L S C P I G I

1321 GAGGACTTCCGGACCTGGTTCATTGACCTGTGGAACAATCCATCATCCCTATCTACAG
E D F R T W F I D L W N N S I I P Y L Q

Fig. 12

74/ 124

1381 GAAGGAGCCAAGGATGGGATCAAGGTTTCATGGACAGAAAGCTGCTTGGGAAGACCCGGTG
E G A K D G I K V H G Q K A A W E D P V

1441 GAATGGGTCCGAGACACTCTTCCCTGGCCGTCGGCCCAACAAGACCAATCAAAGCTCTAC
E W V R D T L P W P S A Q Q D Q S K L Y

1501 CACCTGCCCCCGCCTTCTGTGGGCCCCACAGCACTGCCTCACCCCGGAGGACAGGACA
H L P P P S V G P H S T A S P P E D R T

1561 GTCAAAGACAGCACTCCAAACTCCCTCGACTCAGATCCCCTGATGGCCATGCTACTGAAA
V K D S T P N S L D S D P L M A M L L K

1621 CTCCAAGAAGCTGCCAACTACATTGAGTCACCAGATCGAGAGACTATCCTGGACCCCAAC
L Q E A A N Y I E S P D R E T I L D P N

1681 CTCCAGGCGACACTCTGAGGGCCCGGCAGTCACTGTCACCCTGGAGGGCAGAAGGCTGGC
L Q A T L *

1741 TTCAGCATCATTAGCTCTCCTCTGCCCTCTTCCTTCATAGCTCTGGCTCACCAGCCTCGC

1801 CAAGAGAACAGGAGGGAAGAAGAGGGCAGGAGGAGGGATGGGTTCTCGGTGCTGAACCTT

1861 TGAGAACTTCTACTAGGAATTGGAGGGGGTGAGTTTGAGAACTCCGTGCCCTTAACT

1921 ACATTGCTGGCCTCCTCTTACGACTTAGGAGAAAAGATGATTCTGGTCTTTTCTTCAAG

1981 TTTTGTTCACCTACAAACTCTTGGGCTTCTGAGGAGGGATTGGAAGATATAAACAGA

2041 CAAACAAAACAAACAAACCAACTACAGCAGTTCCAAGCTCGTTCTCACAAACACCTCTG

2101 AGACAGTCACATGTGGGCAAATCTAAGGGAGGCAGGAAGCTCTACAGACTTTCTTGCAAA

2161 CCCTTCCCAGTTCTGTGCGACACTGCCAACAACCTCCCCGCCAGAGACCTGGCCAGAGCCA

2221 AGAAAAGAGAAGCATGTGGTTTAAACAGAAAAACAAAACAAAACAAAAAATATATG

2281 TGTAAATCAACCTGTAGAAGGTAAAAACGGCAATGGAAAAGATGAAGCTGGAAGGAGGGG

2341 CCCAGTTGCCAAGATGGAACGAGAGCTGCCAGATCTTGCCTTCTGGATGACAAGAGGGGA

2401 CATTGCAAGATGGCTGCCAGTCTAAAACGTCACCAGACCACAAGAGTAACATCACAGCCT

2461 TCGAAGAAAGGCCACAAGCTGTCTTTCTGCCCTCTAACTGAACATGCATGAAAAGTCAAT

2521 AAACCCTACTTTTTTAATTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Fig. 12 (cont'd)

T2 Murine cDNA with following intron

```
CCAATAGAACTCCGGATCAAGAGGCAGAATTCCTCAGATAGCATCTCCAGCCTCAACAGC
1  -----+-----+-----+-----+-----+-----+ 60
a  P I E L R I K R Q N S S D S I S S L N S -
ATCACCAGCCATTCCAGCATCGGCAGCAGCAAAGATGCTGATGCCAAGAAGAAAAAGAAG
61  -----+-----+-----+-----+-----+-----+ 120
a  I T S H S S I G S S K D A D A K K K K K -
AAGAGTTGGGTATGTAAAGGCTTGGGGATCGGCCTGTGCTAGGAGTCACTCACCTGTTG
121 -----+-----+-----+-----+-----+-----+ 180
a  K S W
CAGGGAAGTGAACCCCTTTCAGGATCAACAAAGAGGGTCCCTTCTAACAGGATGCCAGTGT
181 -----+-----+-----+-----+-----+-----+ 240
TG TGACATCTGCTGGGGACAAAATTCACTAAGTTCCCATTCCTCTATCCATTGTCTATT
241 -----+-----+-----+-----+-----+-----+ 300
CTCCTTACCACCGCCCTGCACATATACCCCAGCCCCCACCCTCCCTGCATCCTTTATAC
301 -----+-----+-----+-----+-----+-----+ 360
ATGTCTGCTATCCTGGGGCTCTACCTACTGATGAGGTCAAATGTATTTGGCCGTAGAAGG
361 -----+-----+-----+-----+-----+-----+ 420
AGCTGAGAAAATTATTCATGGGTGGGAGAGTGGGGCATGTGGAGAGAATTGTAAAGCCAA
421 -----+-----+-----+-----+-----+-----+ 480
GCAGGGTACTCTAGACGCTCCTGGGGCTGTTGCTTTAGTTTGGGTGAGGAGGCTGTGGAA
481 -----+-----+-----+-----+-----+-----+ 540
CGTCCCCATCGCTCCAAAGCCTGCTTTTGTCTGGTCCAGAGGTGGGTTTGTCTGTGTGG
541 -----+-----+-----+-----+-----+-----+ 600
TATCCCCCTGTAACTCTAAACTGGCTTTGGGTGAGCTTTCTACAATCTGTACGCAGGTG
601 -----+-----+-----+-----+-----+-----+ 660
TAGGGCACTGCCTGACTGACTGAAAGGGAGAGTGACCCAGAGTGAGCGTCTTGTCCCTGT
661 -----+-----+-----+-----+-----+-----+ 720
CCCTGCTGAGGAGGGCTGGCTACAGACTTTGGCCTAGTGCAGACAGGAGCCAGCTGTGTG
721 -----+-----+-----+-----+-----+-----+ 780
GAGAAGCAGCTGTGTGAAATGCATGAGTAGTGTGCTGCTGCTGCTGCTGCTTTCTT
781 -----+-----+-----+-----+-----+-----+ 840
TTCATTGTTTTTTTTTTTTTTTTCTTTCTTTTATTTCTTCAAATGCTGACCTCAAATC
CCTATTTTTTTTTCCAGGTTTATGAGGTAAGAACTCGGTTCTCTCTCGTGCTTTTTCT
901 -----+-----+-----+-----+-----+-----+ 960
TTCCCTTTGCACACCTTCGTGTTTCCAGAGCAAGCACCTCTCTTAAAAAAAAAAAAAAAA
961 -----+-----+-----+-----+-----+-----+ 1020
AAAAA
1021 ----- 1025
```

Fig. 13

76/124

splicing variant 1 (JFC410)

1 AGCGAGTTACTCACGCTTCCCCTCCATCGGAAGCCAGCCAGGCCAAAACCCAGCAAGATA
R V T H A S P P S E A S Q A K T Q Q D M

61 TGCAGTCCAGTCTGGCAGCCAGATATGCAACTCAGTCTAATCACAGTGGGAATTGCAACCA
Q S S L A A R Y A T O S N H S G I A T S

121 GTCAAAAAAGCCTACTAGGCTTCCAGGGCCCTCTAGGGTGCCTGCTGCAGGAAGCAGCA
O K K P T R L P G P S R V P A A G S S S

181 GCAACCTCCACCCACCCTCTAATTTAAATAGGAGAAGTCAGAGCTTTACAGCATTGACA
K V Q G A S N L N R R S Q S F N S I D K

241 AA

bp 1 corresponds to bp 914 of THC

underlined sequence represents further splicing form and is not shown in the THC sequence

Fig. 14

77/124

splicing variant 2

```
1      GGCACCTCACGAGGTCCAGAGCCTGCTCATGAGAACGGGTAGTGTGAGATCTACTCTCTCA
      G T H E V Q S L L M R T G S V R S T L S

61     GAAAGATATACCCCATCATCTCGGCAGGCCAACCAAGAAGAGGGCAAAGAGTGGTTGCGT
      E R Y T P S S R Q A N Q E E G K E W L R

121    TCTCATTCTACTGGAGGGCTTCAGGACACTGGCAACCAG
      S H S T G G L Q D T G N Q
```

bp 1 corresponds to bp 3300 of THC

underlined base pairs -> position of the differentially spliced exon which lacks here but is shown in the THC sequence

Fig. 15

T2-cDNA sequence and T2 protein encoded therein

```

CCGCGGGGCTTCCATCCTTCCTTTGACTGATTTTTAAATTTTAATTTGTATTTTCCCCGC
1  -----+-----+-----+-----+-----+ 60
   R G A S I L P L T D F * I L I C I F P A -

CGCCCCGCCCCCTTTTCCTCCGACCCCGCCCTATCGCTCCCCGGCTTCCCTGCTCTTTCCT
61  -----+-----+-----+-----+-----+ 120
   A P P L F L R P R P I A P R L P C S F L -

TTTTCCCGGCTTCCTTCCTCGCGTTTCTTTCCCTGCGCCCTCGGCTTGCCTCTCTCCCT
121 -----+-----+-----+-----+-----+ 180
   F P G F L P R V S F P C A L G L P L S L -

CCTCCCTCGCTCTCTCCCCCTTCTCTCCCTTCTTCTCGGTTTCTTCCGTCTCTCTCT
181 -----+-----+-----+-----+-----+ 240
   L P R S L P L L S P S S S V S S V L S L -

CCCCCTCCTCCTCCCCCGCCTCCTCCTCCTGCGCTCCCGCCCCCTGCCCTCCCCCGT
241 -----+-----+-----+-----+-----+ 300
   P L L L P R L L L L R S R P L P P P P V -

GCCTGCAGACGCGCGGATCGTCCATGCGCTCCTCGCGGGCAGAATGCTGGGCAGCAGCGT
301 -----+-----+-----+-----+-----+ 360
   P A D A R I V H A L L A G R M L G S S V -

CAAGAGCGTGCAGCCCCGAGGTGGAGCTGAGCAGCGGCGGCGGACGAGGGCGCGGACGA
361 -----+-----+-----+-----+-----+ 420
   K S V Q P E V E L S S G G G D E G A D E -

ACCGCGGGGCGCCGGCAGGAAGGCGGCAGCGGCGGACGGCAGAGGCATGCTGCCAAGCG
421 -----+-----+-----+-----+-----+ 480
   P R G A G R K A A A A D G R G M L P K R -

CGCCAAGGCGCCCGGCGGCGGCGGCGGCATGGCCAAGGCCAGCGGGCTGAGCTGAAGGT
481 -----+-----+-----+-----+-----+ 540
   A K A P G G G G G M A K A S A A E L K V -

CTTCAAGTCCGGCAGCGTGGACAGCCGTGTCCCCGGCGGGCCCGCCGCTCCAACCTGCG
541 -----+-----+-----+-----+-----+ 600
   F K S G S V D S R V P G G P P A S N L R -

CAAGCAGAAGTCACTCACCAACCTCTCTTTTCTACGGAAGCTCCGAGAAAAAGCTGCAGCT
601 -----+-----+-----+-----+-----+ 660
   K Q K S L T N L S F L T D S E K K L Q L -

TTATGAGCCCGAATGGAGCGACGATATGGCCAAGGCGCCCAAAGGCTTAGGCAAGGTGGG
661 -----+-----+-----+-----+-----+ 720
   Y E P E W S D D M A K A P K G L G K V G -

GTCCAAGGGCCGTGAAGCTCCGCTGATGTCCAAGACGCTGTCCAAGTCGGAGCACTCGCT
721 -----+-----+-----+-----+-----+ 780
   S K G R E A P L M S K T L S K S E H S L -

CTTCCAGGCCAAGGGCAGCCCGGCGGCGGCGGCCAAGACCCCCCTGGCTCCGCTCGCGCC
781 -----+-----+-----+-----+-----+ 840
   F Q A K G S P A G G A K T P L A P L A P -

```

Fig. 16

CAACCTGGGAAAGCCGAGCCGGATCCCTCGAGGACCCCTATGCGGAGGTCAAGCCGCTCAG
841 -----+-----+-----+-----+-----+-----+-----+ 900
N L G K P S R I P R G P Y A E V K P L S -

CAAGGCGCCTGAAGCGGCCGTGAGCGAAGATGGCAAATCGGACGACGAGCTGCTCTCCAG
901 -----+-----+-----+-----+-----+-----+-----+ 960
K A P E A A V S E D G K S D D E L L S S -

CAAGGCCAAGGCGCAAAAGAGCTCTGGGCCTGTCCCTCTGCCAAGGGCCAGGAGGAGCG
961 -----+-----+-----+-----+-----+-----+-----+ 1020
K A K A Q K S S G P V P S A K G Q E E R -

CGCCTTCCTCAAGGTGGACCCCGAGCTGGTGGTGACCGTGCTGGGAGACCTGGAGCAGCT
1021 -----+-----+-----+-----+-----+-----+-----+ 1080
A F L K V D P E L V V T V L G D L E Q L -

GCTCTTCAGCCAGATGCTGGACCCAGAGTCCCAGAGAAAGAGGACAGTGCAGAATGTCCT
1081 -----+-----+-----+-----+-----+-----+-----+ 1140
L F S Q M L D P E S Q R K R T V Q N V L -

GGATCTCCGGCAGAACCTGGAAGAGACCATGTCCAGCCTGCGAGGGTCCCAGGTGACTCA
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
D L R Q N L E E T M S S L R G S Q V T H -

CAGCTCCCTGGAGATGACCTGCTACGACAGCGATGATGCCAACCACGCAGCGTGTCCAG
1201 -----+-----+-----+-----+-----+-----+-----+ 1260
S S L E M T C Y D S D D A N P R S V S S -

CCTCTCCAACCGCTCGTACCCTCTGTTCATGGCGCTATGGCCAGTCCAGTCCGCGGCTGCA
1261 -----+-----+-----+-----+-----+-----+-----+ 1320
L S N R S Y P L S W R Y G Q S S P R L Q -

GGCTGGTGACGCGCCCTCTGTGGGTGGGAGCTGCCGCTCGGAGGGGACGCCCGCCTGGTA
1321 -----+-----+-----+-----+-----+-----+-----+ 1380
A G D A P S V G G S C R S E G T P A W Y -

CATGCACGGCGAACGGGCCCCACTACTCCCACACCATGCCCATGCGCAGCCCCAGCAAGCT
1381 -----+-----+-----+-----+-----+-----+-----+ 1440
M H G E R A H Y S H T M P M R S P S K L -

CAGCCATATCTCCCGCCTGGAGCTGGTTCGAATCCCTGGACTCGGATGAGGTGGACCTCAA
1441 -----+-----+-----+-----+-----+-----+-----+ 1500
S H I S R L E L V E S L D S D E V D L K -

GTCCGGCTACATGAGCGACAGTGACCTCATGGGCAAGACCATGACGGAGGATGATGACAT
1501 -----+-----+-----+-----+-----+-----+-----+ 1560
S G Y M S D S D L M G K T M T E D D D I -

CACTACCGGCTGGGATGAAAGCAGCTCCATCAGTAGTGGACTCAGCGATGCCTCAGACAA
1561 -----+-----+-----+-----+-----+-----+-----+ 1620
T T G W D E S S S I S S G L S D A S D N -

TCTCAGTTCAGAAGAATTCAATGCCAGCTCCTCACTCAACTCCCTCCCAAGTACTCCAC
1621 -----+-----+-----+-----+-----+-----+-----+ 1680
L S S E E F N A S S S L N S L P S T P T -

TGCTTCTCGCAGGAAGTCAACAATAGTGCTACGCACAGACTCAGAGAAGCGCTCACTGGC
1681 -----+-----+-----+-----+-----+-----+-----+ 1740
A S R R N S T I V L R T D S E K R S L A -

Fig. 16 (cont'd 1)

80/124

```

AGAAAGTGGGCTGAGCTGGTTTAGTGAATCAGAGGAGAAAGCCCCCTAAAAAACTGGAGTA
1741 -----+-----+-----+-----+-----+-----+-----+ 1800
      E S G L S W F S E S E E K A P K K L E Y -

CGACAGTGGTAGCCTGAAGATGGAACCTGGGACTTCTAAGTGGCGGAGGGAGCGGCCTGA
1801 -----+-----+-----+-----+-----+-----+-----+ 1860
      D S G S L K M E P G T S K W R R E R P E -

GAGCTGTGATGATTCATCCAAGGGTGGAGAACTGAAAAAGCCCCATCAGCCTGGGCCACCC
1861 -----+-----+-----+-----+-----+-----+-----+ 1920
      S C D D S S K G G E L K K P I S L G H P -

TGGTTCCTGAAGAAGGGCAAGACCCACCTGTGGCTGTAACTTCCCCATCACTCACAC
1921 -----+-----+-----+-----+-----+-----+-----+ 1980
      G S L K K G K T P P V A V T S P I T H T -

AGCCCAGAGTGGCCTCAAAGTCGCAGGCAAACCTGAGGGCAAAGCTACAGACAAGGGTAA
1981 -----+-----+-----+-----+-----+-----+-----+ 2040
      A Q S A L K V A G K P E G K A T D K G K -

GCTTGCAGTGAAGAATACTGGGCTCCAACGCTCCTCCTCTGATGCTGGTTCGGGACCGCCT
2041 -----+-----+-----+-----+-----+-----+-----+ 2100
      L A V K N T G L Q R S S S D A G R D R L -

GAGTGATGCTAAGAAGCCCCCTCGGGCATTGCTCGCCCCTCCACTTCGGGATCCTTTGG
2101 -----+-----+-----+-----+-----+-----+-----+ 2160
      S D A K K P P S G I A R P S T S G S F G -

CTACAAGAAGCCTCCTCCTGCCACAGGCACAGCCACTGTCATGCAAAGTGGTGGTTCAGC
2161 -----+-----+-----+-----+-----+-----+-----+ 2220
      Y K K P P P A T G T A T V M Q T G G S A -

CACTCTCAGCAAGATCCAGAAGTCCTCAGGCATCCCTGTCAAGCCAGTAAATGGGCGCAA
2221 -----+-----+-----+-----+-----+-----+-----+ 2280
      T L S K I Q K S S G I P V K P V N G R K -

GACTAGCTTAGATGTTTCCAACAGTGCAGAGCCAGGATTCCTGGCTCCTGGAGCCCGTTC
2281 -----+-----+-----+-----+-----+-----+-----+ 2340
      T S L D V S N S A E P G F L A P G A R S -

TAACATCCAGTACCGCAGCCTGCCCCGGCCAGCCAAGTCAAGTTCTATGAGCGTGACCGG
2341 -----+-----+-----+-----+-----+-----+-----+ 2400
      N I Q Y R S L P R P A K S S S M S V T G -

CGGGCGGGGTGGACCTCGCCCTGTGAGCAGCAGCATTGACCCAGTCTCCTCAGCACCAA
2401 -----+-----+-----+-----+-----+-----+-----+ 2460
      G R G G P R P V S S S I D P S L L S T K -

GCAGGGAGGCCTTACGCCTTCCAGACTGAAGGAGCCTACCAAGGTAGCCAGTGGGCGGAC
2461 -----+-----+-----+-----+-----+-----+-----+ 2520
      Q G G L T P S R L K E P T K V A S G R T -

CACTCCAGCCCCTGTCAATCAGACAGATCGGGAAGGAGAAGGCCAAAGCCAAGGCAGT
2521 -----+-----+-----+-----+-----+-----+-----+ 2580
      T P A P V N Q T D R E K E K A K A K A V -

GGCCTTGGACTCAGACAACATCTCCTTGAAGAGTATTGGCTCCCCAGAAAGTACTCCCAA
2581 -----+-----+-----+-----+-----+-----+-----+ 2640
      A L D S D N I S L K S I G S P E S T P K -

```

Fig. 16 (cont'd 2)

GAACCAAGCAAGCCACCCACAGCCACCAAGCTGGCAGAGCTGCCACCAACCCCTCTCAG
2641 -----+-----+-----+-----+-----+-----+-----+ 2700
N Q A S H P T A T K L A E L P P T P L R -

GGCCACAGCGAAGAGCTTTGTCAAACCACCTCACTAGCCAATCTTGACAAGGTCAACTC
2701 -----+-----+-----+-----+-----+-----+-----+ 2760
A T A K S F V K P P S L A N L D K V N S -

CAACAGTCTGGATCTACCATCATCCAGTGATACCACCCATGCTTCAAAGGTCCCAGATCT
2761 -----+-----+-----+-----+-----+-----+-----+ 2820
N S L D L P S S S D T T H A S K V P D L -

GCATGCTACAAGCTCAGCATCTGGGGGCCCTCTCCCTTCCTGCTTCACCCCCAGTCCGGC
2821 -----+-----+-----+-----+-----+-----+-----+ 2880
H A T S S A S G G P L P S C F T P S P A -

ACCCATCCTCAATATTAACCTCAGCCAGCTTCTCCCAGGGCCTGGAGCTAATGAGTGGTTT
2881 -----+-----+-----+-----+-----+-----+-----+ 2940
P I L N I N S A S F S Q G L E L M S G F -

CAGTGTGCCAAAAGAGACCCGCATGTACCCCAAACCTCTCAGGCCTGCACAGGAGCATGGA
2941 -----+-----+-----+-----+-----+-----+-----+ 3000
S V P K E T R M Y P K L S G L H R S M E -

GTCCCTCCAGATGCCAATGAGCCTCCCCAGTGCCTTCCCCAGCAGTACTCCCGTCCCCAC
3001 -----+-----+-----+-----+-----+-----+-----+ 3060
S L Q M P M S L P S A F P S S T P V P T -

CCCACCTGCTCCCCCTGCTGCTCCACAGAAGAAGAGACGGAAGAGCTGACTTGGAGTGG
3061 -----+-----+-----+-----+-----+-----+-----+ 3120
P P A P P A A P T E E E T E E L T W S G -

AAGCCCCAGAGCTGGGCAACTGGACAGTAATCAGCGGGATCGGAACACTCTTCCCAAGAA
3121 -----+-----+-----+-----+-----+-----+-----+ 3180
S P R A G Q L D S N Q R D R N T L P K K -

AGGGCTCAGGTACCAGCTTCAGTCCCAGGAGGAGACCAAGGAGAGGCGACATTCCCATAC
3181 -----+-----+-----+-----+-----+-----+-----+ 3240
G L R Y Q L Q S Q E E T K E R R H S H T -

CATTGGTGGGCTGCCTGAATCCGATGACCAGTCAGAGCTGCCTTCTCCCCCTGCACTTCC
3241 -----+-----+-----+-----+-----+-----+-----+ 3300
I G G L P E S D D Q S E L P S P P A L P -

CATGTCTCTGAGTGCAAAGGGCCAACTTACCAACATAGTGAGTCCCACTGCGGCCACCAC
3301 -----+-----+-----+-----+-----+-----+-----+ 3360
M S L S A K G Q L T N I V S P T A A T T -

*GCCAAGAATCACCCGCTCCAACAGCATCCCCACCCACGAGGCGGCCTTCGAGCTGTACAG
3361 -----+-----+-----+-----+-----+-----+-----+ 3420
P R I T R S N S I P T H E A A F E L Y S -

CGGCTCCCAAATGGGGAGCACCCCTGTCCCTGGCCGAGAGACCCAAGGGAATGATTCGGTC
3421 -----+-----+-----+-----+-----+-----+-----+ 3480
G S Q M G S T L S L A E R P K G M I R S -

AGGATCCTTCCGAGACCCACGGACGATGTTACGGCTCAGTGCTGTCCCTGGCCTCCAG
3481 -----+-----+-----+-----+-----+-----+-----+ 3540
G S F R D P T D D V H G S V L S L A S S -

Fig. 16 (cont'd 3)

TGCCTCCTCCACCTACTCCTCAGCTGAGGAGAGGATGCAATCTGAGCAAATCCGGAAGCT
 3541 -----+-----+-----+-----+-----+-----+-----+ 3600
 A S S T Y S S A E E R M Q S E Q I R K L -
 TCGTAGGGAACTGGAATCATCCCAGGAAAAAGTGGCCACCTTGACGTCTCAGCTTTCTGC
 3601 -----+-----+-----+-----+-----+-----+-----+ 3660
 R R E L E S S Q E K V A T L T S Q L S A -
 CAATGCTAATCTGGTGGCTGCTTTTGAGCAGAGCCTGGTGAATATGACATCCCGCCTGCG
 3661 -----+-----+-----+-----+-----+-----+-----+ 3720
 N A N L V A A F E Q S L V N M T S R L R -
 ACACCTGGCAGAGACGGCCGAGGAGAAGGACACTGAGCTGCTGGATTTGCGAGAAACCAT
 3721 -----+-----+-----+-----+-----+-----+-----+ 3780
 H L A E T A E E K D T E L L D L R E T I -
 AGACTTTCTGAAGAAAAAGAACTCTGAGGCCAGGCAGTCATTTCAGGGAGCCCTTAATGC
 3781 -----+-----+-----+-----+-----+-----+-----+ 3840
 D F L K K K N S E A Q A V I Q G A L N A -
 CTCAGAAACCACACCCAAAGAACTTCGGATCAAGAGACAAAACCTCCTCAGATAGCATCTC
 3841 -----+-----+-----+-----+-----+-----+-----+ 3900
 S E T T P K E L R I K R Q N S S D S I S -
 AAGCCTCAACAGCATCACTAGCCATTCCAGCATCGGCAGCAGCAAGGATGCTGATGCGAA
 3901 -----+-----+-----+-----+-----+-----+-----+ 3960
 S L N S I T S H S S I G S S K D A D A K -
 AAAGAAGAAAAAAGAGTTGGCTTCGAAGTTCCTTCAACAAAGCGTTTCAGTATAAAAAA
 3961 -----+-----+-----+-----+-----+-----+-----+ 4020
 K K K K K S W L R S S F N K A F S I K K -
 GGGGCCCAAGTCAGCTTCCTCATACTCGGATATAGAGGAGATTGCTACACCCGACTCTTC
 4021 -----+-----+-----+-----+-----+-----+-----+ 4080
 G P K S A S S Y S D I E E I A T P D S S -
 AGCCCCCTCATCCCCAAACTACAGCATGGTTCTACAGAGACTGCTTCACCCCTCCATCAA
 4081 -----+-----+-----+-----+-----+-----+-----+ 4140
 A P S S P K L Q H G S T E T A S P S I K -
 GTCCTCCACCTCGTCCTCCGTGGGCACTGATGTCACCGAGGGCCCTGCTCAGCCAGCCCC
 4141 -----+-----+-----+-----+-----+-----+-----+ 4200
 S S T S S S V G T D V T E G P A H P A P -
 CCACACTAGGCTGTTCCATGCAAATGAGGAGGAGGAGCCAGAGAAGAAGGAGGTATCGGA
 4201 -----+-----+-----+-----+-----+-----+-----+ 4260
 H T R L F H A N E E E E P E K K E V S E -
 GCTGCGCTCTGAGCTATGGGAGAAGGAAATGAAGCTTACAGACATCCGCTTGAGGCCCT
 4261 -----+-----+-----+-----+-----+-----+-----+ 4320
 L R S E L W E K E M K L T D I R L E A L -
 CAACTCTGCCCCACCAACTGGATCAGCTTCGGGAGACCATGCACAACATGCAGTTGGAGGT
 4321 -----+-----+-----+-----+-----+-----+-----+ 4380
 N S A H Q L D Q L R E T M H N M Q L E V -
 GGACCTGCTGGAAGCAGAGAATGACCGACTGAAGGTAGCCCCAGGCCCTCATCAGGCTC
 4381 -----+-----+-----+-----+-----+-----+-----+ 4440
 D L L E A E N D R L K V A P G P S S Q S -

83/124

CACTCCAGGGCAGGTCCCTGGATCATCTGCATTATCTTCCCCACGCCGCTCCCTAGGCCT
 4441 -----+-----+-----+-----+-----+-----+-----+ 4500
 T P G Q V P G S S A L S S P R R S L G L -
 GGCACCTACCCATTCTTTCGGCCCCAGTCTTGCAGACACAGACCTGTACCCCATGGATGG
 4501 -----+-----+-----+-----+-----+-----+-----+ 4560
 A L T H S F G P S L A D T D L S P M D G -
 CATCAGTACTTGTGGTCCAAAGGAGGAAGTGACCCTCCGGGTGGTGGTGAGGATGCCCCC
 4561 -----+-----+-----+-----+-----+-----+-----+ 4620
 I S T C G P K E E V T L R V V V R M P P -
 GCAGCACATCATCAAAGGGGACTTGAAGCAGCAGGAATTCTTCCTGGGCTGTAGCAAGGT
 4621 -----+-----+-----+-----+-----+-----+-----+ 4680
 Q H I I K G D L K Q Q E F F L G C S K V -
 CAGTGGAAAAGTTGACTGGAAGATGCTGGATGAAGCTGTTTTCCAAGTGTTCAAGGACTA
 4681 -----+-----+-----+-----+-----+-----+-----+ 4740
 S G K V D W K M L D E A V F Q V F K D Y -
 TATTTCTAAAATGGACCCAGCCTCTACCCTGGGACTAAGCACTGAGTCCATCCATGGCTA
 4741 -----+-----+-----+-----+-----+-----+-----+ 4800
 I S K M D P A S T L G L S T E S I H G Y -
 CAGCATCAGCCACGTGAAACGAGTGTTGGATGCAGAGCCCCCGAGATGCCTCCTTGCCG
 4801 -----+-----+-----+-----+-----+-----+-----+ 4860
 S I S H V K R V L D A E P P E M P P C R -
 TCGAGGTGTCAATAACATATCAGTCTCCCTCAAAGGTCTGAAGGAGAAATGCGTCGACAG
 4861 -----+-----+-----+-----+-----+-----+-----+ 4920
 R G V N N I S V S L K G L K E K C V D S -
 CCTGGTGTTCGAGACGCTGATCCCCAAGCCGATGATGCAGCACTACATAAGCCTCCTGCT
 4921 -----+-----+-----+-----+-----+-----+-----+ 4980
 L V F E T L I P K P M M Q H Y I S L L L -
 GAAGCACCGGCGCCTCGTCTCTCGGGCCCCAGCGGCACGGGCAAGACCTACCTGACCAA
 4981 -----+-----+-----+-----+-----+-----+-----+ 5040
 K H R R L V L S G P S G T G K T Y L T N -
 TCGCTTGGCCGAGTACCTGGTGGAGCGCTCTGGCCGTGAGGTCACAGAGGGCATCGTCAG
 5041 -----+-----+-----+-----+-----+-----+-----+ 5100
 R L A E Y L V E R S G R E V T E G I V S -
 CACCTTCAACATGCACCAGCAGTCTTGCAAGGATCTGCAACTGTATCTTTCCAACCTAGC
 5101 -----+-----+-----+-----+-----+-----+-----+ 5160
 T F N M H Q Q S C K D L Q L Y L S N L A -
 CAACCAGATAGACCGGGAACAGGAATTGGGGATGTGCCCTGGTGATTCTATTGGATGA
 5161 -----+-----+-----+-----+-----+-----+-----+ 5220
 N Q I D R E T G I G D V P L V I L L D D -
 CCTGAGTGAAGCAGGCTCCATCAGTGAGTTGGTCAATGGGGCCCTCACCTGCAAGTATCA
 5221 -----+-----+-----+-----+-----+-----+-----+ 5280
 L S E A G S I S E L V N G A L T C K Y H -
 TAAATGTCCTATATTATAGGTACCACCAATCAGCCTGTAAAAATGACACCCAACCATGG
 5281 -----+-----+-----+-----+-----+-----+-----+ 5340
 K C P Y I I G T T N Q P V K M T P N H G -

Fig. 16 (cont'd 5)

```

CTTTCAGCTTGAGCTTCAGGATGTTGACCTTCTCCAACAACGTGGAGCCAGCCAATGGCTT
5341 -----+-----+-----+-----+-----+-----+-----+ 5400
      F H L S F R M L T F S N N V E P A N G F -

CCTGGTTCGTTACCTGAGGAGGAAGCTGGTAGAGTCAGACAGCGACATCAATGCCAACAA
5401 -----+-----+-----+-----+-----+-----+-----+ 5460
      L V R Y L R R K L V E S D S D I N A N K -

GGAAGAGCTGCTTCGGGTGCTCGACTGGGTACCCAAGCTGTGGTATCATCTCCACACCTT
5461 -----+-----+-----+-----+-----+-----+-----+ 5520
      E E L L R V L D W V P K L W Y H L H T F -

CCTTGAGAAGCACAGCACCTCAGACTTCCTCATCGGCCCTTGCTTCTTTCTGTCGTGTCC
5521 -----+-----+-----+-----+-----+-----+-----+ 5580
      L E K H S T S D F L I G P C F F L S C P -

CATTGGCATTGAGGACTTCCGGACCTGGTTTCATTGACCTGTGGAACAACCTCTATCATTC
5581 -----+-----+-----+-----+-----+-----+-----+ 5640
      I G I E D F R T W F I D L W N N S I I P -

CTATCTACAGGAAGGAGCCAAGGATGGGATAAAGGTCCATGGACAGAAAGCTGCTTGGGA
5641 -----+-----+-----+-----+-----+-----+-----+ 5700
      Y L Q E G A K D G I K V H G Q K A A W E -

GGACCCAGTGGAATGGGTCCGGGACACACTTCCCTGGCCATCAGCCCAACAAGACCAATC
5701 -----+-----+-----+-----+-----+-----+-----+ 5760
      D P V E W V R D T L P W P S A Q Q D Q S -

AAAGCTGTACCACCTGCCCCACCCACCGTGGGCCCTCACAGCATTGCCTCACCTCCCGA
5761 -----+-----+-----+-----+-----+-----+-----+ 5820
      K L Y H L P P P T V G P H S I A S P P E -

GGATAGGACAGTCAAAGACAGCACCCCAAGTTCTCTGGACTCAGATCCTCTGATGGCCAT
5821 -----+-----+-----+-----+-----+-----+-----+ 5880
      D R T V K D S T P S S L D S D P L M A M -

GCTGCTGAAACTTCAAGAAGCTGCCAACTACATTGAGTCTCCAGATCGAGAAACCATCCT
5881 -----+-----+-----+-----+-----+-----+-----+ 5940
      L L K L Q E A A N Y I E S P D R E T I L -

GGACCCCAACCTTCAGGCAACACTTTAAGGGTTCGGCAATCACTGTACCCCCGGACAGC
5941 -----+-----+-----+-----+-----+-----+-----+ 6000
      D P N L Q A T L * -

AGAACGCTGGCATCAGCTATCTTAGCTCCTCCTCTCCCCTCTCCTCTTTTCAGAGCACTGG
6001 -----+-----+-----+-----+-----+-----+-----+ 6060

CTCTCCAGCCCCAGGAGGAGAACAGGAGGGAGGAGGAGATGAAAGAGGAGGGACAGGTTC
6061 -----+-----+-----+-----+-----+-----+-----+ 6120

TTGGTGCTGTACCTTTGAGAACTTCCTAGGAAGGAATGGTGGGGTGGCGTTTGGGAACTT
6121 -----+-----+-----+-----+-----+-----+-----+ 6180

GTGCCCCCTAAACACATTTACTGGCCTCCTCTAATGACTTTGGGGAAAAGATGATTCTGG
6181 -----+-----+-----+-----+-----+-----+-----+ 6240

GTCTTTCCCTTGACTTCTTGTTTCAATTACAACTCCTGGGCTTTCTGGGGAGGGGTTC
6241 -----+-----+-----+-----+-----+-----+-----+ 6300

```

Fig. 16 (cont'd 6)

GAAAACATCAAAACACTGCAGCAGTTCCTAAATGATTCTCACAAGCAACCCTGAGAGAGA
6301 -----+-----+-----+-----+-----+-----+ 6360
CAGTCTTGTGAGGGAGATCTGGGGGAGGCAGGAAGCTCCTCAGATTTTCTCACAGACCCT
6361 -----+-----+-----+-----+-----+-----+ 6420
TCCCAATTCCATCACCCTGCGCAACAACCTCCTCCCCCAGAGATCTGGCTGGAGCCCAGAA
6421 -----+-----+-----+-----+-----+-----+ 6480
AAAGAAGCATGTGGTTTAAAAAATGTTTAAATCAATCTGTAAAAGGTAAAAATGAAAAAC
6481 -----+-----+-----+-----+-----+-----+ 6540
AAAAACAAGCAAACAAACAAAAACAATGGAAAAGATGAAGCTGGAGAGAGAGGAACCAG
6541 -----+-----+-----+-----+-----+-----+ 6600
TTGCCAAGGTAGAGAGCTGCCCCGCTCCTGCCCTCTGGATGACATAGGGGACATCAACAAG
6601 -----+-----+-----+-----+-----+-----+ 6660
ACGGCTGCCAACCTGAGAAGTCACCAAACCACAAAAATAACCTTACAGCCTTCAGGGAAA
6661 -----+-----+-----+-----+-----+-----+ 6720
GACTACCAGCTCTGTCTTTCTACCCTCTAATTTAACAATGCATAAGAGTCAATAAACCTT
6721 -----+-----+-----+-----+-----+-----+ 6780
ACTTTTTTTAAAAAAG
6781 -----+-----+-----+-----+-----+ 6805

Fig. 16 (cont'd 7)

T3-cDNA sequence and T3 protein
encoded therein (protein isoform 1)

```
CAACCAGCCAGAACGCCTGAACTCGCAGGTGCTGCAGGGGCTGCAGGAGCCAGCGGGGGA
1  -----+-----+-----+-----+-----+-----+ 60
  N Q P E R L N S Q V L Q G L Q E P A G E -

GGGGCTCCCGCTGCGGAAGAGCGGCTCGGTGGAACCGGGTTCGATACCCAGATCTACAC
61  -----+-----+-----+-----+-----+-----+ 120
  G L P L R K S G S V E N G F D T Q I Y T -

AGACTGGGCCAATCATTACCTAGCCAAATCCGGCCACAAGCGTCTCATCAGGGATCTCCA
121 -----+-----+-----+-----+-----+-----+ 180
  D W A N H Y L A K S G H K R L I R D L Q -

GCAAGATGTGACAGATGGCGTCCTCCTGGCCCAGATTATCCAGGTTGTGGCAAATGAAAA
181 -----+-----+-----+-----+-----+-----+ 240
  Q D V T D G V L L A Q I I Q V V A N E K -

GATTGAAGACATCAATGGCTGTCCGAAGAACAGATCCCAAATGATTGAAAACATAGATGC
241 -----+-----+-----+-----+-----+-----+ 300
  I E D I N G C P K N R S Q M I E N I D A -

CTGCTTGAATTTCTGCGCAGCTAAGGGAATAAACATCCAGGGGCTGTCTGCAGAAGAGAT
301 -----+-----+-----+-----+-----+-----+ 360
  C L N F L A A K G I N I Q G L S A E E I -

CAGGAATGGAAACCTCAAGGCCATTCTAGGCCTCTTCTTCAGCCTCTCCCGATACAAGCA
361 -----+-----+-----+-----+-----+-----+ 420
  R N G N L K A I L G L F F S L S R Y K Q -

GCAGCAGCAGCAGCCCCAGAAGCAGCACCTCTCCTCACCTCTGCCGCCCCGCGTATCCCA
421 -----+-----+-----+-----+-----+-----+ 480
  Q Q Q Q P Q K Q H L S S P L P P A V S Q -

GGTGGCCGGGGCCCCCTCCCAGTGCCAGGCTGGCACCCCTCAGCAGCAGGTGCCAGTCAC
481 -----+-----+-----+-----+-----+-----+ 540
  V A G A P S Q C Q A G T P Q Q Q V P V T -

TCCCCAAGCCCCGTGCCAGCCTCACCAGCCAGCGCCACATCAGCAGTCAAAGCACAAGC
541 -----+-----+-----+-----+-----+-----+ 600
  P Q A P C Q P H Q P A P H Q Q S K A Q A -

TGAAATGCAGTCCAGACTTCCAGGTCTACCGCGAGGGTATCCGCTGCAGGCAGCGAGGC
601 -----+-----+-----+-----+-----+-----+ 660
  E M Q S R L P G P T A R V S A A G S E A -

CAAAACACGCGGAGGGTCAACTACTGCTAACAACCGACGCAGCCAGAGCTTTAACAACCTA
661 -----+-----+-----+-----+-----+-----+ 720
  K T R G G S T T A N N R R S Q S F N N Y -

TGATAAATCCAAACCAGTCACCTCCCCACCCCCACCGCCAAGCAGCCACGAGAAAGAGCC
721 -----+-----+-----+-----+-----+-----+ 780
  D K S K P V T S P P P P P S S H E K E P -

TTTGGCAAGTTTCAGCCTCCTCCCACCCCGGAATGAGTGACAATGCACCTGCTTCCTTGGA
781 -----+-----+-----+-----+-----+-----+ 840
  L A S S A S S H P G M S D N A P A S L E -
```

Fig. 17

GAGCGGCAGCAGCTCCACCCCTACTAATTGCAGTACCTCCTCGGCCATCCCGCAGCCCGG
841 -----+-----+-----+-----+-----+-----+-----+-----+ 900
S G S S S T P T N C S T S S A I P Q P G -
TGCAGCCACCAAGCCTTGGCGCAGCAAATCCCTCAGCGTGAAGCACAGTGCCACGGTATC
901 -----+-----+-----+-----+-----+-----+-----+ 960
A A T K P W R S K S L S V K H S A T V S -
CATGCTCTCGGTCAAGCCTCCTGGGCCTGAGGCCCCCAGGCCCACACCTGAAGCCATGAA
961 -----+-----+-----+-----+-----+-----+-----+ 1020
M L S V K P P G P E A P R P T P E A M K -
GCCGGCCCCCAACAATCAGAAGTCCATGCTGGAAAAGCTGAAACTTTTCAACAGTAAAGG
1021 -----+-----+-----+-----+-----+-----+-----+ 1080
P A P N N Q K S M L E K L K L F N S K G -
GGGCTCAAAGGCAGGTGAGGGGCCGGGGTCCCGGGACACAAGCTGTGAGCGGCTGGAGAC
1081 -----+-----+-----+-----+-----+-----+-----+ 1140
G S K A G E G P G S R D T S C E R L E T -
TCTGCCCAGCTTCGAAGAGAGCGAGGAGCTGGAGGCCGCCAGTCGCATGCTCACCACCGT
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
L P S F E E S E E L E A A S R M L T T V -
GGGCCCTGCTTCCAGCAGCCCCAAGATTGCACTCAAGGGCATTGCCCAGAGGACTTTTAG
1201 -----+-----+-----+-----+-----+-----+-----+ 1260
G P A S S S P K I A L K G I A Q R T F S -
CCGGGCACTGACCAACAAGAAGAGTTCTCTGAAAGGCAATGAGAAAGAGAAGGAGAAACA
1261 -----+-----+-----+-----+-----+-----+-----+ 1320
R A L T N K K S S L K G N E K E K E K Q -
ACAGCGGGAGAAGGATAAGGAGAAAAGCAAGGACCTTGCCAAGAGAGCCTCTGTGACGGA
1321 -----+-----+-----+-----+-----+-----+-----+ 1380
Q R E K D K E K S K D L A K R A S V T E -
GAGGCTGGACCTCAAGGAGGAGCCAAAAGAAGACCCCACTGGAGCAGCTGTGCCCCGAGAT
1381 -----+-----+-----+-----+-----+-----+-----+ 1440
R L D L K E E P K E D P S G A A V P E M -
GCCAAAAAAGTCCTCCAAGATTGCCAGCTTCATCCCCAAAGGGGGGAAGCTCAACAGTGC
1441 -----+-----+-----+-----+-----+-----+-----+ 1500
P K K S S K I A S F I P K G G K L N S A -
CAAGAAGGAGCCCATGGCCCCCTTCCCACAGTGGAATACCAAACCAGGAATGAAGAGCAT
1501 -----+-----+-----+-----+-----+-----+-----+ 1560
K K E P M A P S H S G I P K P G M K S M -
GCCCCGGAAATCCCCAAGTGCCCCAGCGCCTTCCAAGGAAGGGGAGCGGAGCCGGAGTGG
1561 -----+-----+-----+-----+-----+-----+-----+ 1620
P G K S P S A P A P S K E G E R S R S G -
GAAGCTGAGCTCAGGACTCCCCCAGCAGAAGCCCCAGCTGGACGGCAGACACTCCAGTTC
1621 -----+-----+-----+-----+-----+-----+-----+ 1680
K L S S G L P Q Q K P Q L D G R H S S S -
CTCTTCCAGCCTGGCGTCCTCAGAAGGAAAAGGCCAGGAGGGACCACCCTGAACCACAG
1681 -----+-----+-----+-----+-----+-----+-----+ 1740
S S S L A S S E G K G P G G T T L N H S -

Fig. 17 (cont'd 1)

88/124

CATCAGCAGCCAGACTGTCAGTGGGTCTGTCTCGGGACCACCCAGACCACAGGAAGCAATAC
 1741 -----+-----+-----+-----+-----+-----+-----+ 1800
 I S S Q T V S G S V G T T Q T T G S N T -
 CGTCAGTGTTCAGCTACCTCAGCCCCAGCAGCAATACAACCATCCCAACACTGCCACGGT
 1801 -----+-----+-----+-----+-----+-----+-----+ 1860
 V S V Q L P Q P Q Q Q Y N H P N T A T V -
 TGCACCTTTCTGTACAGGTCTCAGACGGACACTGAAGGGAATGTTACTGCCGAGTCAAG
 1861 -----+-----+-----+-----+-----+-----+-----+ 1920
 A P F L Y R S Q T D T E G N V T A E S S -
 CTCAACAGGTGTGAGCGTGGAGCCCAGCCACTTCACCAAGACTGGACAGCCTGCTCTGGA
 1921 -----+-----+-----+-----+-----+-----+-----+ 1980
 S T G V S V E P S H F T K T G Q P A L E -
 AGAACTCACTGGGGAAGATCCTGAGGCTCGGCGGCTGCGGACAGTGAAGAACATCGCTGA
 1981 -----+-----+-----+-----+-----+-----+-----+ 2040
 E L T G E D P E A R R L R T V K N I A D -
 TCTGCGGCAGAAATTTGGAGGAAACCATGTCCAGTTTAAGGGGAACCTCAGGTTACACACAG
 2041 -----+-----+-----+-----+-----+-----+-----+ 2100
 L R Q N L E E T M S S L R G T Q V T H S -
 CACATTGGAAACCACGTTTGACACCAATGTCACCACGGAGATGAGTGGCCGTAGCATACT
 2101 -----+-----+-----+-----+-----+-----+-----+ 2160
 T L E T T F D T N V T T E M S G R S I L -
 CAGCTTGACAGGGAGGCCCACACCTCTGTCTGGAGACTGGGCCAGTCCAGCCCTCGGCT
 2161 -----+-----+-----+-----+-----+-----+-----+ 2220
 S L T G R P T P L S W R L G Q S S P R L -
 CCAAGCAGGAGACGCCCCCTCAATGGGCAATGGGTATCCCCCTCGAGCCAACGCCAGCAG
 2221 -----+-----+-----+-----+-----+-----+-----+ 2280
 Q A G D A P S M G N G Y P P R A N A S R -
 GTTCATCAACACTGAGTCAGGTCGCTATGTGTACTCCGCCCCCTCTGAGAAGGCAGCTGGC
 2281 -----+-----+-----+-----+-----+-----+-----+ 2340
 F I N T E S G R Y V Y S A P L R R Q L A -
 CTCCCCGGGCGAGTAGTGTCTGCCACGTGGACGTCTCAGACAAGGCAGGAGATGAGATGGA
 2341 -----+-----+-----+-----+-----+-----+-----+ 2400
 S R G S S V C H V D V S D K A G D E M D -
 CCTGGAAGGCATCAGCATGGACGCCCCCGGCTACATGAGCGATGGGGATGTTCTGAGCAA
 2401 -----+-----+-----+-----+-----+-----+-----+ 2460
 L E G I S M D A P G Y M S D G D V L S K -
 GAACATCCGGACCGATGACATTACAAGCGGATACATGACTGATGGTGGACTTGGCCTCTA
 2461 -----+-----+-----+-----+-----+-----+-----+ 2520
 N I R T D D I T S G Y M T D G G L G L Y -
 TACCCGTCGCCTGAACCGGCTCCCTGATGGGATGGCTGTGGTACGGGAGACCCTGCAACG
 2521 -----+-----+-----+-----+-----+-----+-----+ 2580
 T R R L N R L P D G M A V V R E T L Q R -
 AAATACCTCCCTGGGCCTCGGAGACGCTGACAGCTGGGACGACAGCAGCTCCGTCAGCAG
 2581 -----+-----+-----+-----+-----+-----+-----+ 2640
 N T S L G L G D A D S W D D S S S V S S -

Fig. 17 (cont'd 2)

CGGCATCAGCGACACCATAGACAACCTCAGCACTGATGACATCAACACCAGCTCCTCCAT
 2641 -----+-----+-----+-----+-----+-----+-----+ 2700
 G I S D T I D N L S T D D I N T S S S I -
 CAGCTCTTATGCCAACACACCTGCCTCCTCTCGAAAAAACCTGGATGTGCAGACTGATGC
 2701 -----+-----+-----+-----+-----+-----+-----+ 2760
 S S Y A N T P A S S R K N L D V Q T D A -
 TGAGAAGCACTCACAGGTGGAGAGGAATTCCCTGTGGTCTGGTGTATGATGTCAAGAAATC
 2761 -----+-----+-----+-----+-----+-----+-----+ 2820
 E K H S Q V E R N S L W S G D D V K K S -
 AGACGGAGGCTCAGACAGCGGCATAAAAAATGGAGCCAGGTTCCAAGTGGAGGCGGAATCC
 2821 -----+-----+-----+-----+-----+-----+-----+ 2880
 D G G S D S G I K M E P G S K W R R N P -
 TTCTGATGTGTCTGACGAGTCCGACAAAAGCACGTCGGGCAAGAAGAATCCTGTCTATCTC
 2881 -----+-----+-----+-----+-----+-----+-----+ 2940
 S D V S D E S D K S T S G K K N P V I S -
 CCAGACAGGCTCATGGCGGCGAGGCATGACAGCTCAGGTGGGCATCACCATGCCAAGGAC
 2941 -----+-----+-----+-----+-----+-----+-----+ 3000
 Q T G S W R R G M T A Q V G I T M P R T -
 GAAGGCTTCAGCCCCGGCAGGCGCACTGAAGACCCCAGGAAGTGGAAAAACAGACGACGC
 3001 -----+-----+-----+-----+-----+-----+-----+ 3060
 K A S A P A G A L K T P G T G K T D D A -
 AAAGGTGTCTGAGAAAGGAAGGCTTTCTCCTAAAGCCTCCCAGGTGAAGCGCTCCCCATC
 3061 -----+-----+-----+-----+-----+-----+-----+ 3120
 K V S E K G R L S P K A S Q V K R S P S -
 AGATGCAGGCCGGAGCAGTGGTGACGAATCCAAAAAGCCCCTCCCCAGCAGCTCTAGGAC
 3121 -----+-----+-----+-----+-----+-----+-----+ 3180
 D A G R S S G D E S K K P L P S S S R T -
 ACCTACTGCCAATGCCAACAGCTTTGGGTTCAAGAAGCAGAGTGGTTCGCCACCGGCCCT
 3181 -----+-----+-----+-----+-----+-----+-----+ 3240
 P T A N A N S F G F K K Q S G S A T G L -
 GGCCATGATCACAGCCAGCGGGGTGACTGTCACCAGCAGGTCAGCCACACTGGGCAAAAT
 3241 -----+-----+-----+-----+-----+-----+-----+ 3300
 A M I T A S G V T V T S R S A T L G K I -
 CCCAAAGTCATCTGCACTCGTCAGTCGGTCTGCTGGTTCGGAAGTCAAGTATGGATGGGGC
 3301 -----+-----+-----+-----+-----+-----+-----+ 3360
 P K S S A L V S R S A G R K S S M D G A -
 TCAGAATCAGGATGACGGGTATCTAGCCCTAAGCTCCCGGACAAACCTTCAGTACCGGAG
 3361 -----+-----+-----+-----+-----+-----+-----+ 3420
 Q N Q D D G Y L A L S S R T N L Q Y R S -
 TTTGCCGAGGCCAGTAAGTCCAACAGCCGGAACGGGGCTGGGAACAGGTCTAGCACCAG
 3421 -----+-----+-----+-----+-----+-----+-----+ 3480
 L P R P S K S N S R N G A G N R S S T S -
 CAGCATAGATTCCAACATTAGCAGCAAGTCCGCAGGCCTGCCAGTGCCCCAAACTGAGGGA
 3481 -----+-----+-----+-----+-----+-----+-----+ 3540
 S I D S N I S S K S A G L P V P K L E -

Fig. 17 (cont'd 3)

```

GCCTTCCAAAACAGCCCTAGGCAGCTCTCTACCAGGTCTGGTCAACCAAACAGACAAGGA
3541 -----+-----+-----+-----+-----+-----+-----+ 3600
    P S K T A L G S S L P G L V N Q T D K E -
GAAAGGCATCTCATCAGACAACGAGAGTGTGGCTTCTGTAACTCGGTGAAAGTGAATCC
3601 -----+-----+-----+-----+-----+-----+-----+ 3660
    K G I S S D N E S V A S C N S V K V N P -
GGCAGCCCAGCCTGTGTCCAGTCCGGCTCAGACCAGTCTCCAGCCTGGAGCCAAGTACCC
3661 -----+-----+-----+-----+-----+-----+-----+ 3720
    A A Q P V S S P A Q T S L Q P G A K Y P -
AGATGTGGCCTCTCCACACTCCGCAGACTCTTTGGTGGGAAGCCTACCAAGCAAGTGCC
3721 -----+-----+-----+-----+-----+-----+-----+ 3780
    D V A S P T L R R L F G G K P T K Q V P -
CATCGCCACAGCTGAAAACATGAAAAATTCGGTGGTCATCTCCAATCCTCATGCCACCAT
3781 -----+-----+-----+-----+-----+-----+-----+ 3840
    I A T A E N M K N S V V I S N P H A T M -
GACTCAGCAAGGTAACCTAGACTCCCCGTGAGGAGTGGCGTCCTGAGCAGTGGGAGCAG
3841 -----+-----+-----+-----+-----+-----+-----+ 3900
    T Q Q G N L D S P S G S G V L S S G S S -
CAGTCTCTCTACAGCAAGAATGTGGACCTCAACCAGTCTCCGCTAGCCTCCAGCCCCAG
3901 -----+-----+-----+-----+-----+-----+-----+ 3960
    S P L Y S K N V D L N Q S P L A S S P S -
CTCAGCCCAGTCCGCCCCCTTCCAACAGCCTCACCTGGGGGCACCAACGCCAGCAGCTCCTC
3961 -----+-----+-----+-----+-----+-----+-----+ 4020
    S A H S A P S N S L T W G T N A S S S S -
CGCAGTTAGCAAGGATGGCCTGGGCTTTTCAGTCTGTCAGCAGCCTCCACACCAGCTGTGA
4021 -----+-----+-----+-----+-----+-----+-----+ 4080
    A V S K D G L G F Q S V S S L H T S C E -
GTCCATCGACATCTCCCTCAGCAGTGGAGGGGTCCCCAGCCACAATTCTTCCACTGGCCT
4081 -----+-----+-----+-----+-----+-----+-----+ 4140
    S I D I S L S S G G V P S H N S S T G L -
CATCGCCTCCTCCAAGGACGACTCCTTGACTCCCTTTGTCAGAACTAACAGTGTGAAGAC
4141 -----+-----+-----+-----+-----+-----+-----+ 4200
    I A S S K D D S L T P F V R T N S V K T -
CACACTGTCAGAAAGCCCTCTCTCTTCCCCTGCTGCTAGCCCTAAGTTCTGCAGAAGTAC
4201 -----+-----+-----+-----+-----+-----+-----+ 4260
    T L S E S P L S S P A A S P K F C R S T -
TCTGCCCAGGAAACAGGACAGTGACCCGCACCTTGATAGGAACACTTTGCCTAAGAAAGG
4261 -----+-----+-----+-----+-----+-----+-----+ 4320
    L P R K Q D S D P H L D R N T L P K K G -
ACTCAGGTATACTCCACCTCCCAGCTTCGCACGCAAGAAGATGCAAAGAATGGTTACG
4321 -----+-----+-----+-----+-----+-----+-----+ 4380
    L R Y T P T S Q L R T Q E D A K E W L R -
GTCCCATTCTGCAGGAGGCCTTCAGGACACCGCTGCCAATTCCCCCTTTTCTCTGGCTC
4381 -----+-----+-----+-----+-----+-----+-----+ 4440
    S H S A G G L Q D T A A N S P F S S G S -

```

Fig. 17 (cont'd 4)

CAGCGTGACTTCTCCCTCCGGAACAAGATTCAACTTTTCCCAGCTTGCGAGTCCCACCAC
 4441 -----+-----+-----+-----+-----+ 4500
 S V T S P S G T R F N F S Q L A S P T T -
 TGTCACCCAGATGAGCTTGTCCAACCCGACCATGCTGAGGACTCACAGCCTCTCCAATGC
 4501 -----+-----+-----+-----+-----+ 4560
 V T Q M S L S N P T M L R T H S L S N A -
 TGATGGGCAGTATGATCCATACTGACAGCCGCTTCCGGAATAGCTCCATGTCCCTGGA
 4561 -----+-----+-----+-----+-----+ 4620
 D G Q Y D P Y T D S R F R N S S M S L D -
 TGAGAAGAGCAGAACCATGAGCCGTTTCAGGCTCATTCCGGGATGGGTTTGAAGAAGTTCA
 4621 -----+-----+-----+-----+-----+ 4680
 E K S R T M S R S G S F R D G F E E V H -
 TGGATCCTCACTCTCCTTGGTTTCCAGCACATCGTCAGTTTATTCTACACCAGAAGAAAA
 4681 -----+-----+-----+-----+-----+ 4740
 G S S L S L V S S T S S V Y S T P E E K -
 ATGCCAGTCAGAGATTGCAAGCTGCGGCGGGAAGTGGATGCCTCCCAGGAGAAAGTTTC
 4741 -----+-----+-----+-----+-----+ 4800
 C Q S E I R K L R R E L D A S Q E K V S -
 AGCTTTGACCACCCAGCTGACAGCAAATGCTCACCTTGTGGCTGCCTTTGAACAGAGTCT
 4801 -----+-----+-----+-----+-----+ 4860
 A L T T Q L T A N A H L V A A F E Q S L -
 TGGTAACATGACAATCAGGCTCCAGAGTCTGACCATGACAGCTGAGCAGAAGGATTTCAGA
 4861 -----+-----+-----+-----+-----+ 4920
 G N M T I R L Q S L T M T A E Q K D S E -
 ACTGAATGAGTTAAGAAAAACCATTGAGCTGCTAAAGAAACAGAACGCAGCTGCCCAGGC
 4921 -----+-----+-----+-----+-----+ 4980
 L N E L R K T I E L L K K Q N A A A Q A -
 TGCCATTAATGGAGTAATTAACACACCTGAGCTCAACTGCAAAGGAAACGGCACTGCCCCA
 4981 -----+-----+-----+-----+-----+ 5040
 A I N G V I N T P E L N C K G N G T A Q -
 GTCTGCAGACCTCCGCATCCGCAGGCAGCACTCCTCAGACAGCGTCTCCAGCATCAACAG
 5041 -----+-----+-----+-----+-----+ 5100
 S A D L R I R R Q H S S D S V S S I N S -
 TGCCACCAGCCACTCCAGTGTGGGCAGCAACATAGAGAGTGAAGTCAAAGAAGAAGAAGAG
 5101 -----+-----+-----+-----+-----+ 5160
 A T S H S S V G S N I E S D S K K K K R -
 GAAGAACTGGGTCAATGAGTTACGCAGCTCCTTCAAGCAAGCTTTCGGAAGAAGAAGTC
 5161 -----+-----+-----+-----+-----+ 5220
 K N W V N E L R S S F K Q A F G K K K S -
 CCCAAAATCTGCGTCCTCTCATTGAGATATTGAGGAGATGACGGATTCTTCTTTGCCTTC
 5221 -----+-----+-----+-----+-----+ 5280
 P K S A S S H S D I E E M T D S S L P S -
 CTCACCAAAGTTACCACACAATGGGTCCACAGGTTCCACCCCACTGCTGAGGAATTCTCA
 5281 -----+-----+-----+-----+-----+ 5340
 S P K L P H N G S T G S T P L L R N S H -

CTCCAACCTCTCTAATTTTCAGAATGCATGGATAGTGAAGCTGAGACCGTCATGCAGCTCCG
5341 -----+-----+-----+-----+-----+ 5400
S N S L I S E C M D S E A E T V M Q L R -
AAATGAGTTAAGAGACAAGGAGATGAAGCTGACAGATATCCGCTTAGAAGCTCTCAGTTC
5401 -----+-----+-----+-----+-----+ 5460
N E L R D K E M K L T D I R L E A L S S -
TGCCCCACCAGCTGGACCAGCTCCGGGAGGCCATGAACAGGATGCAGAGTGAAATAGAGAA
5461 -----+-----+-----+-----+-----+ 5520
A H Q L D Q L R E A M N R M Q S E I E K -
GCTGAAAGCTGAGAATGATCGGCTGAAGTCAGAGTCTCAAGGCAGTGGCTGCAGCCGGGC
5521 -----+-----+-----+-----+-----+ 5580
L K A E N D R L K S E S Q G S G C S R A -
TCCTTCCCAAGTGTCCATCTCTGCCTCCCCGAGGCAGTCCATGGGCCTCTCCCAGCACAG
5581 -----+-----+-----+-----+-----+ 5640
P S Q V S I S A S P R Q S M G L S Q H S -
CTTGAACCTCACTGAGTCAACCAGCCTGGACATGTTGCTGGATGACACTGGTGAATGCTC
5641 -----+-----+-----+-----+-----+ 5700
L N L T E S T S L D M L L D D T G E C S -
GGCTCGGAAGGAAGGAGGCAGGCATGTTAAGATAGTTGTCAGCTTTCAGGAGGAAATGAA
5701 -----+-----+-----+-----+-----+ 5760
A R K E G G R H V K I V V S F Q E E M K -
GTGGAAGGAGGATTCCAGACCACATCTCTTTCTTATTGGCTGCATTGGAGTTAGTGGCAA
5761 -----+-----+-----+-----+-----+ 5820
W K E D S R P H L F L I G C I G V S G K -
GACGAAGTGGGATGTGCTCGATGGGGTGGTTAGACGGCTGTTCAAAGAATACATCATTCA
5821 -----+-----+-----+-----+-----+ 5880
T K W D V L D G V V R R L F K E Y I I H -
TGTCGACCCAGTGAGTCAGCTAGGGCTGAATTACAGACAGCGTTCTTGGCTACAGCATTGG
5881 -----+-----+-----+-----+-----+ 5940
V D P V S Q L G L N S D S V L G Y S I G -
AGAAATCAAGCGCAGCAACACTTCCGAAACACCGGAGCTGCTTCCTTGTGGCTATCTGGT
5941 -----+-----+-----+-----+-----+ 6000
E I K R S N T S E T P E L L P C G Y L V -
TGGAGAGAACACGACCATCTCAGTGAAGTGTGAAAGGGCTCGCAGAAAACAGCCTGGACTC
6001 -----+-----+-----+-----+-----+ 6060
G E N T T I S V T V K G L A E N S L D S -
ACTGGTGTGTTGAGTCCTTGATTCCCAAGCCCATCCTGCAGCGCTACGTCTCCCTCCTGAT
6061 -----+-----+-----+-----+-----+ 6120
L V F E S L I P K P I L Q R Y V S L L I -
AGAGCACCGTCGGATCATTCTCTGCCCCAGCGGCACTGGGAAAACCTACCTGGCCAA
6121 -----+-----+-----+-----+-----+ 6180
E H R R I I L S G P S G T G K T Y L A N -
CCGGCTGTCTGAGTATATAGTGCTTCGAGAGGGACGGGAGTTGACAGACGGGGTTATCGC
6181 -----+-----+-----+-----+-----+ 6240
R L S E Y I V L R E G R E L T D G V I A -

Fig. 17 (cont'd 6)

CACCTTTAACGTGACCATAAGTCCAGCAAGGAATTGCGCCAGTACCTGTCCAACCTTGC
 6241 -----+-----+-----+-----+-----+ 6300
 T F N V D H K S S K E L R Q Y L S N L A -
 TGACCAGTGC AACAGTGAGAACAATGCTGTGGACATGCCCCCTCGTCATCATCTGGACAA
 6301 -----+-----+-----+-----+-----+ 6360
 D Q C N S E N N A V D M P L V I I L D N -
 CCTACACCACGTGAGCTCTCTGGGCGAGATCTTCAATGGGCTGCTCAACTGCAAGTACCA
 6361 -----+-----+-----+-----+-----+ 6420
 GGATGTGGTGC ACTCGAGAGACCCGCTCTAGAAGTTACCCGACGAGTTGACGTTTCATGGT
 L H H V S S L G E I F N G L L N C K Y H -
 CAAATGCCCTTACATAATTGGCACAATGAACCAGGCTACCTCTTCGACTCCCAACCTGCA
 6421 -----+-----+-----+-----+-----+ 6480
 GTTTACGGGAATGTATTAACCGTGTTACTTGGTCCGATGGAGAAGCTGAGGGTTGGACGT
 K C P Y I I G T M N Q A T S S T P N L Q -
 GCTTCACCATAACTTCAGATGGGTGCTTTGTGCCAACCACACGGAGCCTGTGAAGGGTTT
 6481 -----+-----+-----+-----+-----+ 6540
 CGAAGTGGTATTGAAGTCTACCCACGAAACACGGTTGGTGTGCCTCGGACACTTCCCAA
 L H H N F R W V L C A N H T E P V K G F -
 CCTTGGCCGATTCTGAGGAGGAAGCTCATGGAAACAGAGATCAGTGGGCGGGTGCGCAA
 6541 -----+-----+-----+-----+-----+ 6600
 GGAACCGGCTAAGGACTCCTCCTTCGAGTACCTTTGTCTCTAGTCACCCGCCCACGCGTT
 L G R F L R R K L M E T E I S G R V R N -
 TATGGAGCTGGTAAAAATCATTGACTGGATTCCCAAGGTCTGGCATCACCTCAACCGCTT
 6601 -----+-----+-----+-----+-----+ 6660
 ATACCTCGACCATTTTTAGTAAGTACCTAAGGGTTCCAGACCGTAGTGGAGTTGGCGAA
 M E L V K I I D W I P K V W H H L N R F -
 CCTGGAGGCTCACAGTTCCTCGGACGTCACCATCGGCCCCCGGCTCTTCCTGTCATGCCC
 6661 -----+-----+-----+-----+-----+ 6720
 L E A H S S S D V T I G P R L F L S C P -
 CATCGATGTGGACGGCTCGAGAGTGTGGTTACCGACTTGTGGAAGTATTCCATTATCCC
 6721 -----+-----+-----+-----+-----+ 6780
 I D V D G S R V W F T D L W N Y S I I P -
 CTATCTCCTGGAAGCCGTCAGAGAAGGACTCCAGCTCTATGGAAGGCGCGCCCCCTGGGA
 6781 -----+-----+-----+-----+-----+ 6840
 Y L L E A V R E G L Q L Y G R R A P W E -
 GGATCCTGCCAAGTGGGTGATGGACACATATCCATGGGCAGCCAGCCACAACAGCACGA
 6841 -----+-----+-----+-----+-----+ 6900
 D P A K W V M D T Y P W A A S P Q Q H E -
 GTGGCCTCCCCTGCTGCAGTTACGGCCTGAGGATGTGGGCTTCGACGGCTACTCCATGCC
 6901 -----+-----+-----+-----+-----+ 6960
 W P P L L Q L R P E D V G F D G Y S M P -
 TCGGGAGGGATCGACAAGCAAGCAGATGCCCCCAGTGATGCTGAAGGTGACCCGCTGAT
 6961 -----+-----+-----+-----+-----+ 7020
 R E G S T S K Q M P P S D A E G D P L M -
 GAACATGCTGATGAGGCTGCAGGAGGCAGCCAACTACTCCAGCCCCCAGAGCTATGACAG
 7021 -----+-----+-----+-----+-----+ 7080
 N M L M R L Q E A A N Y S S P Q S Y D S -

CGACTCCAACAGCAACAGCCATCACGATGACATCTTGGACTCCTCTTTGGAGTCCACTCT
7081 -----+-----+-----+-----+-----+ 7140
D S N S N S H H D D I L D S S L E S T L -
GTGACAGGGGCGGAGCCAGCGCCCTCCTCTTCTCCTCACCGCATTCCACCTGCATCC
7141 -----+-----+-----+-----+-----+ 7200
*
CCCACATCACCTGAAGATGACTTCCTGAGCCAGCCCCAGCCACAGCCTTAGAGCTGCG
7201 -----+-----+-----+-----+-----+ 7260
GGAACACCGAGACCCCCGTCCTTCAGCCTCGACCTGGGTGCAGGCATCCCGGGCCAGCT
7261 -----+-----+-----+-----+-----+ 7320
GCCTGCGGACCGCTTCCTTCCACAGCGAGAACTGCACTACCTTCTGTTGTACTTTAATTA
7321 -----+-----+-----+-----+-----+ 7380
TTGTTTTGCCTTGTTGCTGTGACCTCCCTAAGACACTGAAGATACTTCTCGGGAAAGGAT
7381 -----+-----+-----+-----+-----+ 7440
CATCGCCGTTGAAATGAAAAGAGAGACAGAGAGAGAAAAAAGAGAACCCACATGAA
7441 -----+-----+-----+-----+-----+ 7500
GCTCTGAAACCAAACAGCATCCTGCCATGAGCTTCCCAGAGACAGAAGAGACTGGAGCAA
7501 -----+-----+-----+-----+-----+ 7560
AGTCGGAACACAGAGAAGCACGGCTTCCCCTCAGCACAGACCCTCCAGACTGGGTCTCA
7561 -----+-----+-----+-----+-----+ 7620
GAGCCGTGCCACCCACCCTCCCACACAGCCGGCCACAGGGAGAACTGGTGCTAACCAGGG
7621 -----+-----+-----+-----+-----+ 7680
TGCTTGCTTTGGTCACGTTCAACGCACTACAGAGCTACGACACAGGGGAACCTTAGGAGC
7681 -----+-----+-----+-----+-----+ 7740
AAATAAACCGTGCTTTCATGTTTTTTAAAAAAAAAAAAAAAAAAAA
7741 -----+-----+-----+-----+-----+ 7783

Fig. 17 (cont'd 8)

**T3-cDNA sequence and T3 protein encoded
therein (isoform 2)**

```
AGCAGGGAGAGGGGAGGGAGTGTGCCGTCTCTTCTGCAAGGGCAGTGCCCCAGCCTCAGC
1  -----+-----+-----+-----+-----+ 60
  S R E R G G S V P S L L Q G Q C P S L S -

CACACTTCTGATCTGCAGTCCAACAGACCTTTCTAGCATGCCAAAGAGAACCTGGGGGTG
61  -----+-----+-----+-----+-----+ 120
  H T S D L Q S N R P F * H A K E N L G V -

CCAGGGGGTCCTCAGAGCTCACACTGCACTTGTGGCACCCACAGCGAGTAGCCATCCGTG
121 -----+-----+-----+-----+-----+ 180
  P G G P Q S S H C T C G T H S E * P S V -

AGCCGAGGAAACTGTACACAGATCTACACAGACTGGGCCAATCATTACCTAGCCAAATCC
181 -----+-----+-----+-----+-----+ 240
  S R G N C T Q I Y T D W A N H Y L A K S -

GGCCACAAGCGTCTCATCAAGGATCTCCAGCAAGATGTGACAGATGGCGTCCTCCTGGCC
241 -----+-----+-----+-----+-----+ 300
  G H K R L I K D L Q Q D V T D G V L L A -

CAGATTATCCAGGTTGTGGCAAATGAAAAGATTGAAGACATCAATGGCTGTCCGAAGAAC
301 -----+-----+-----+-----+-----+ 360
  Q I I Q V V A N E K I E D I N G C P K N -

AGATCCCCAAATGATTGAAAACATAGATGCCTGCTTGAATTTCTGGCAGCTAAGGGAATA
361 -----+-----+-----+-----+-----+ 420
  R S Q M I E N I D A C L N F L A A K G I -

AACATCCAGGGGCTGTCTGCAGAAGAGATCAGGAATGGAAACCTCAAGGCCATTCTAGGC
421 -----+-----+-----+-----+-----+ 480
  N I Q G L S A E E I R N G N L K A I L G -

CTCTTCTTCAGCCTCTCCCGATACAAGCAGCAGCAGCAGCAGCCCCAGAAGCAGCACCTC
481 -----+-----+-----+-----+-----+ 540
  L F F S L S R Y K Q Q Q Q Q P Q K Q H L -

TCCTCACCTCTGCGCGCCGCGTATCCAGGTGGCCGGGGCCCCCTCCAGTGCCAGGCT
541 -----+-----+-----+-----+-----+ 600
  S S P L P P A V S Q V A G A P S Q C Q A -

GGCACCCCTCAGCAGCAGGTGCCAGTCACTCCCCAAGCCCCGTGCCAGCCTCACCAGCCA
601 -----+-----+-----+-----+-----+ 660
  G T P Q Q Q V P V T P Q A P C Q P H Q P -
```

Fig. 18

96/124
T3 murine cDNA

```

ATGAGAAGAGCCGAACAATGAGTCGGTCAGGCTCCTTCCGGGATGGGTTTGAGGAAGTTC
1  -----+-----+-----+-----+-----+ 60
   E K S R T M S R S G S F R D G F E E V H -

ATGGATCCTCCCTGTCCTTGGTTTCCAGCACATCCTCCATCTACTCCACGCCAGAAGAAA
61  -----+-----+-----+-----+-----+ 120
   G S S L S L V S S T S S I Y S T P E E K -

AATGCCAGTCAGAGATTCGAAAGCTGAGGCGAGAACTGGATGCCTCCCAGGAAAAGGTGT
121 -----+-----+-----+-----+-----+ 180
   C Q S E I R K L R R E L D A S Q E K V S -

AATGCCAGTCAGAGATTCGAAAGCTGAGGCGAGACGTGGATGCCTCCCAGGAAAAGGTGT
121 -----+-----+-----+-----+-----+ 180
   C Q S E I R K L R R D V D A S Q E K V S -

CTGCGCTGACTACCCAGCTGACTGCAAATGCTCACCTTGTGGCAGCCTTCGAGCAGAGTC
181 -----+-----+-----+-----+-----+ 240
   A L T T Q L T A N A H L V A A F E Q S L -

TGGGAAACATGACCATCAGGCTACAGAGTTTAACTATGACCGCTGAGCAGAAGGATTTCAG
241 -----+-----+-----+-----+-----+ 300
   G N M T I R L Q S L T M T A E Q K D S E -

AACTGAACGAGTTAAGAAAAACCATCGAGCTGCTGAAGAAACAGAATGCAGCTGCCCAGG
301 -----+-----+-----+-----+-----+ 360
   L N E L R K T I E L L K K Q N A A A Q A -

CTGCCATTAATGGAGTGATTAACACGCCAGAGCTCAACTGCAAAGGAAATGGCAGTGCCA
361 -----+-----+-----+-----+-----+ 420
   A I N G V I N T P E L N C K G N G S A R -

GGCTACAGACCTACGCATCCGCAGCAACACTCCTCCGACAGTGTCTCCAGTATCAATAGC
421 -----+-----+-----+-----+-----+ 480
   L Q T Y A S A A T L L R Q C L Q Y Q * R -

GCCACCAGCCACTCAAGTGTG
481 -----+-----+ 501
   H Q P L K C -

```

Fig. 19

T2
CAGCCTCTCCAACCGCTCGTA
S L S N R S

AGCCTCTCCAACCGCTCGTAC
S L S N R S Y

GECTCTCCAACCGCTCGTACC
S L S N R S Y

CCTCTCCAACCGCTCGTACCC
L S N R S Y

CTCTCCAACCGCTCGTACCCT
L S N R S Y P

TCTCCAACCGCTCGTACCCTC
L S N R S Y P

CTCCAACCGCTCGTACCCTCT
S N R S Y P

TCCAACCGCTCGTACCCTCTG
S N R S Y P L

CCAACCGCTCGTACCCTCTGT
S N R S Y P L

CAACCGCTCGTACCCTCTGTC
N R S Y P L

AACCGCTCGTACCCTCTGTCA
N R S Y P L S

ACCGCTCGTACCCTCTGTCA
N R S Y P L S

CCGCTCGTACCCTCTGTCA
R S Y P L S

CGCTCGTACCCTCTGTCA
R S Y P L S W

GCTCGTACCCTCTGTCA
S Y P L S W

CTCGTACCCTCTGTCA
S Y P L S W

TCGTACCCTCTGTCA
S Y P L S W R

CGTACCCTCTGTCA
S Y P L S W R

GTACCCTCTGTCA
Y P L S W R

TACCCTCTGTCA
Y P L S W R Y

ACCCTCTGTCA
Y P L S W R Y

T2 97/124
CCTCTCCACCTACTCCTCAC
A S S T Y S S

CTCTCCACCTACTCCTCACA
S S T Y S S

TCCTCCACCTACTCCTCACA
S S T Y S S Q

CCTCCACCTACTCCTCACAAA
S S T Y S S Q

CTCCACCTACTCCTCACAAAT
S T Y S S Q

TCCACCTACTCCTCACAAATC
S T Y S S Q I

CCACCTACTCCTCACAAATCC
S T Y S S Q I

CACCTACTCCTCACAAATCCG
T Y S S Q I

ACCTACTCCTCACAAATCCGG
T Y S S Q I R

CCTACTCCTCACAAATCCGGA
T Y S S Q I R

CTACTCCTCACAAATCCGGAA
Y S S Q I R

TACTCCTCACAAATCCGGAAG
Y S S Q I R K

ACTCCTCACAAATCCGGAAGC
Y S S Q I R K

CTCCTCACAAATCCGGAAGCT
S S Q I R K

TCCTCACAAATCCGGAAGCTT
S S Q I R K L

CCTCACAAATCCGGAAGCTTC
S S Q I R K L

CTCACAAATCCGGAAGCTTCG
S Q I R K L

TCACAAATCCGGAAGCTTCGT
S Q I R K L R

CACAAATCCGGAAGCTTCGTA
S Q I R K L R

ACAAATCCGGAAGCTTCGTAG
Q I R K L R

097914549 097914549 097914549

AGAAGAAAAAAGAGTTGGC
K K K K S W

GAAGAAAAAAGAGTTGGCT
K K K K S W

AAAGAAAAAAGAGTTGGCTT
K K K K S W L

AGAAAAAAGAGTTGGCTTC
K K K K S W L

GAAAAAAAGAGTTGGCTTCG
K K K S W L

AAAAAAGAGTTGGCTTCGA
K K K S W L R

AAAAAAGAGTTGGCTTCGAA
K K K S W L R

AAAAAAGAGTTGGCTTCGAAG
K K S W L R

AAAAAAGAGTTGGCTTCGAAGT
K K S W L R S

AAAAAAGAGTTGGCTTCGAAGTT
K K S W L R S

AAAGAGTTGGCTTCGAAGTTTC
K S W L R S

AAGAGTTGGCTTCGAAGTTTCC
K S W L R S S

AGAGTTGGCTTCGAAGTTCTCT
K S W L R S S

GAGTTGGCTTCGAAGTTCTCTT
S W L R S S

AGTTGGCTTCGAAGTTCTCTTC
S W L R S S F

GTGGCTTCGAAGTTCTCTTCA
S W L R S S F

TTGGCTTCGAAGTTCTCTTCAA
W L R S S F

TGGCTTCGAAGTTCTCTTCAAC
W L R S S F N

GGCTTCGAAGTTCTCTTCAACA
W L R S S F N

GCTTCGAAGTTCTCTTCAACAA
L R S S F N

Fig. 20

T2	T2	98/124
CTCCATCAAGTCCTCCACCTC S I K S S T	AGTTGGAGGTGGACCTGCTGG L E V D L L	ATGACACCCCAACCATGGCTTT M T P N H G F
TCCATCAAGTCCTCCACCTCG S I K S S T S	GTTGGAGGTGGACCTGCTGGA L E V D L L	TGACACCCCAACCATGGCTTTC M T P N H G F
CCATCAAGTCCTCCACCTCGT S I K S S T S	TTGGAGGTGGACCTGCTGGAA L E V D L L E	GACACCCCAACCATGGCTTTCA T P N H G F
CATCAAGTCCTCCACCTCGTC I K S S T S	TGGAGGTGGACCTGCTGGAAG L E V D L L E	ACACCCCAACCATGGCTTTTCA T P N H G F H
ATCAAGTCCTCCACCTCGTCC I K S S T S S	GGAGGTGGACCTGCTGGAAGC E V D L L E	CACCCCAACCATGGCTTTTCACT T P N H G F H
TCAAGTCCTCCACCTCGTCTT I K S S T S S	GAGGTGGACCTGCTGGAAGCA E V D L L E A	ACCCCAACCATGGCTTTTCACTT P N H G F H
CAAGTCCTCCACCTCGTCTCT K S S T S S	AGGTGGACCTGCTGGAAGCAG E V D L L E A	CCCAACCATGGCTTTTCACTTG P N H G F H L
AAGTCCTCCACCTCGTCTCTC K S S T S S S	GGTGGACCTGCTGGAAGCAGA V D L L E A	CCAACCATGGCTTTTCACTTGA P N H G F H L
AGTCCTCCACCTCGTCTCTCCG K S S T S S S	GTGGACCTGCTGGAAGCAGAG V D L L E A E	CAACCATGGCTTTTCACTTGAG N H G F H L
GTCCTCCACCTCGTCTCTCCGT S S T S S S	TGGACCTGCTGGAAGCAGAGA V D L L E A E	AACCATGGCTTTTCACTTGAGC N H G F H L S
TCCTCCACCTCGTCTCTCCGTG S S T S S S V	GGACCTGCTGGAAGCAGAGAA D L L E A E	ACCATGGCTTTTCACTTGAGCT N H G F H L S
CCTCCACCTCGTCTCTCCGTGG S S T S S S V	GACCTGCTGGAAGCAGAGAAT D L L E A E N	CCATGGCTTTTCACTTGAGCTT H G F H L S
CTCCACCTCGTCTCTCCGTGGG S T S S S V	ACCTGCTGGAAGCAGAGAATG D L L E A E N	CATGGCTTTTCACTTGAGCTTC H G F H L S F
TCCACCTCGTCTCTCCGTGGGC S T S S S V G	CCTGCTGGAAGCAGAGAATGA L L E A E N	ATGGCTTTTCACTTGAGCTTCA H G F H L S F
CCACCTCGTCTCTCCGTGGGCA T S S S S V G	CTGCTGGAAGCAGAGAATGAC L L E A E N D	TGGCTTTTCACTTGAGCTTCAG G F H L S F
CACCTCGTCTCTCCGTGGGCAC T S S S S V G	TGCTGGAAGCAGAGAATGACC L L E A E N D	GGCTTTTCACTTGAGCTTCAGG G F H L S F R
ACCTCGTCTCTCCGTGGGCACT T S S S S V G G	GCTGGAAGCAGAGAATGACCG L E A E N D	GCTTTCACTTGAGCTTCAGGA G F H L S F R
CCTCGTCTCTCCGTGGGCACTG T S S S S V G G	CTGGAAGCAGAGAATGACCGA L E A E N D R	CTTTCACTTGAGCTTCAGGAT F H L S F R
CTCGTCTCTCCGTGGGCACTGA S S S V G G	TGGAAGCAGAGAATGACCGAC L E A E N D R	TTTCACTTGAGCTTCAGGATG F H L S F R M
TCGTCTCTCCGTGGGCACTGAT S S S V G G T	GGAAGCAGAGAATGACCGACT E A E N D R	TTCACTTGAGCTTCAGGATGT F H L S F R M
CGTCTCTCCGTGGGCACTGATG S S S V G G T	GAAGCAGAGAATGACCGACTG E A E N D R L	TCACTTGAGCTTCAGGATGTT H L S F R M

Fig. 20 (cont'd 1)

TAAAAGGTAAAAATGAAAAAC
AAAAGGTAAAAATGAAAAACA
AAAGGTAAAAATGAAAAACAA
AAGGTAAAAATGAAAAACAAA
AGGTAAAAATGAAAAACAAA
GGTAAAAATGAAAAACAAA
GTAAAAATGAAAAACAAAAC
TAAAAATGAAAAACAAAACA
AAAAATGAAAAACAAAACA
AAAATGAAAAACAAAACAAG
AAATGAAAAACAAAACAAGC
AATGAAAAACAAAACAAGCA
ATGAAAAACAAAACAAGCAA
TGAAAAACAAAACAAGCAAA
GAAAAACAAAACAAGCAAA
AAAAACAAAACAAGCAAA
AAACAAAACAAGCAAA
AAACAAAACAAGCAAA
AACAAAACAAGCAAA
ACAAAACAAGCAAA

99/124

T2

CTCTAATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
CCTCTAATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
CTCTAATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
TCTAATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
CTAATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
TAATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
AATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
ATTTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
TTTAAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
TTAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
TAACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
AACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
ACAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
CAATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
AATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
ATGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
TGCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
GCATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
CATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
ATAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA
TAAGAGTCAATAAACCCCTACTTTTTTAAAAAAA

Fig. 20 (cont'd 2)

T3

ACTGGGCCAATCATTACCTAG
W A N H Y L
CTGGGCCAATCATTACCTAGC
W A N H Y L
TGGGCCAATCATTACCTAGCC
W A N H Y L A
GGGCCAATCATTACCTAGCCA
W A N H Y L A
GGCCAATCATTACCTAGCCAA
A N H Y L A
GCCAATCATTACCTAGCCAAA
A N H Y L A K
CCAATCATTACCTAGCCAAAT
A N H Y L A K
CAATCATTACCTAGCCAAATC
N H Y L A K
AATCATTACCTAGCCAAATCC
N H Y L A K S
ATCATTACCTAGCCAAATCCG
N H Y L A K S
TCATTACCTAGCCAAATCCGG
H Y L A K S
CATTACCTAGCCAAATCCGGC
H Y L A K S G
ATTACCTAGCCAAATCCGGCC
H Y L A K S G
TTACCTAGCCAAATCCGGCCA
Y L A K S G
TACCTAGCCAAATCCGGCCAC
Y L A K S G H
ACCTAGCCAAATCCGGCCACA
Y L A K S G H
CCTAGCCAAATCCGGCCACAA
L A K S G H
CTAGCCAAATCCGGCCACAAG
L A K S G H K
TAGCCAAATCCGGCCACAAGC
L A K S G H K
AGCCAAATCCGGCCACAAGCG
A K S G H K
GCCAAATCCGGCCACAAGCGT
A K S G H K R

T3

CGGCCACAAGCGTCTCATCAG
G H K R L I
GGCCACAAGCGTCTCATCAGG
G H K R L I R
GCCACAAGCGTCTCATCAGGG
G H K R L I R
CCACAAGCGTCTCATCAGGGA
H K R L I R
CACAAGCGTCTCATCAGGGAT
H K R L I R D
ACAAGCGTCTCATCAGGGATC
H K R L I R D
CAAGCGTCTCATCAGGGATCT
K R L I R D
AAGCGTCTCATCAGGGATCTC
K R L I R D L
AGCGTCTCATCAGGGATCTCC
K R L I R D L
GCGTCTCATCAGGGATCTCCA
R L I R D L
CGTCTCATCAGGGATCTCCAG
R L I R D L Q
GTCTCATCAGGGATCTCCAGC
R L I R D L Q
TCTCATCAGGGATCTCCAGCA
L I R D L Q
CTCATCAGGGATCTCCAGCAA
L I R D L Q Q
TCATCAGGGATCTCCAGCAAG
L I R D L Q Q
CATCAGGGATCTCCAGCAAGA
I R D L Q Q
ATCAGGGATCTCCAGCAAGAT
I R D L Q Q D
TCAGGGATCTCCAGCAAGATG
I R D L Q Q D
CAGGGATCTCCAGCAAGATGT
R D L Q Q D
AGGGATCTCCAGCAAGATGTG
R D L Q Q D V
GGGATCTCCAGCAAGATGTGA
R D L Q Q D V

100/124

09/914549

CTGAAATGCAGTCCAGACTTC
E M Q S R L
TGAAATGCAGTCCAGACTTCC
E M Q S R L
GAAATGCAGTCCAGACTTCCA
E M Q S R L P
AAATGCAGTCCAGACTTCCAG
E M Q S R L P
AATGCAGTCCAGACTTCCAGG
M Q S R L P
ATGCAGTCCAGACTTCCAGGT
M Q S R L P G
TGCAGTCCAGACTTCCAGGTG
M Q S R L P G
GCAGTCCAGACTTCCAGGTCC
Q S R L P G
CAGTCCAGACTTCCAGGTCTC
Q S R L P G P
AGTCCAGACTTCCAGGTCTCA
Q S R L P G P
GTCCAGACTTCCAGGTCTTAC
S R L P G P
TCCAGACTTCCAGGTCTTACC
S R L P G P T
CCAGACTTCCAGGTCTTACCG
S R L P G P T
CAGACTTCCAGGTCTTACCGC
R L P G P T
AGACTTCCAGGTCTTACCGCG
R L P G P T A
GACTTCCAGGTCTTACCGCGA
R L P G P T A
ACTTCCAGGTCTTACCGCGAG
L P G P T A
CTTCCAGGTCTTACCGCGAGG
L P G P T A R
TTCCAGGTCTTACCGCGAGGG
L P G P T A R
TCCAGGTCTTACCGCGAGGGT
P G P T A R
CCAGGTCTTACCGCGAGGGTA
P G P T A R V

Fig. 20 (cont'd 3)

T3

CGGGGCAGTAGTGTCTGCCAC
R G S S V C H

GGGGCAGTAGTGTCTGCCACG
R G S S V C H

GGGCAGTAGTGTCTGCCACGT
G S S V C H

GGCAGTAGTGTCTGCCACGTG
G S S V C H V

GCAGTAGTGTCTGCCACGTGG
G S S V C H V

CAGTAGTGTCTGCCACGTGGA
S S V C H V

AGTAGTGTCTGCCACGTGGAC
S S V C H V D

GTAGTGTCTGCCACGTGGACG
S S V C H V D

AGTAGTGTCTGCCACGTGGACGT
S V C H V D

AGTGTCTGCCACGTGGACGTG
S V C H V D V

GTGTCTGCCACGTGGACGTCT
S V C H V D V

TGTCTGCCACGTGGACGTCTC
V C H V D V

GTCTGCCACGTGGACGTCTCA
V C H V D V S

TCTGCCACGTGGACGTCTCAG
V C H V D V S

CTGCCACGTGGACGTCTCAGA
C H V D V S

TGCCACGTGGACGTCTCAGAC
C H V D V S D

GCCACGTGGACGTCTCAGACA
C H V D V S D

CCACGTGGACGTCTCAGACAA
H V D V S D

CACGTGGACGTCTCAGACAAG
H V D V S D K

ACGTGGACGTCTCAGACAAGG
H V D V S D K

CGTGGACGTCTCAGACAAGGC
V D V S D K

T3

101/124

TCACCATGCCAAGGACGAAGG
T M P R T K

CACCATGCCAAGGACGAAGGC
T M P R T K

ACCATGCCAAGGACGAAGGCT
T M P R T K A

CCATGCCAAGGACGAAGGCTT
T M P R T K A

CATGCCAAGGACGAAGGCTTC
M P R T K A

ATGCCAAGGACGAAGGCTTCA
M P R T K A S

TGCCAAGGACGAAGGCTTCAG
M P R T K A S

GCCAAGGACGAAGGCTTCAGC
P R T K A S

CCAAGGACGAAGGCTTCAGCC
P R T K A S A

CAAGGACGAAGGCTTCAGCCC
P R T K A S A

AAGGACGAAGGCTTCAGCCCC
R T K A S A

AGGACGAAGGCTTCAGCCCCG
R T K A S A P

GGACGAAGGCTTCAGCCCCGG
R T K A S A P

GACGAAGGCTTCAGCCCCGGC
T K A S A P

ACGAAGGCTTCAGCCCCGGCA
T K A S A P A

CGAAGGCTTCAGCCCCGGCAG
T K A S A P A

GAAGGCTTCAGCCCCGGCAGG
K A S A P A

AAGGCTTCAGCCCCGGCAGGC
K A S A P A G

AGGCTTCAGCCCCGGCAGGCG
K A S A P A G

GGCTTCAGCCCCGGCAGGCGC
A S A P A G

GCTTCAGCCCCGGCAGGCGCA
A S A P A G A

09914549 09/914549

AGAAGCAGAGTGGTTCCGCCA
K Q S G S A

GAAGCAGAGTGGTTCCGCCAC
K Q S G S A

AAGCAGAGTGGTTCCGCCACC
K Q S G S A T

AGCAGAGTGGTTCCGCCACCAG
K Q S G S A T

GCAGAGTGGTTCCGCCACCAGG
Q S G S A T

CAGAGTGGTTCCGCCACCAGGC
Q S G S A T G

AGAGTGGTTCCGCCACCAGGCC
Q S G S A T G

GAGTGGTTCCGCCACCAGGCCT
S G S A T G

AGTGGTTCCGCCACCAGGCCTG
S G S A T G L

GTGGTTCCGCCACCAGGCCTGG
S G S A T G L

TGGTTCCGCCACCAGGCCTGGC
G S A T G L

GGTTCCGCCACCAGGCCTGGCC
G S A T G L A

GTTCGCCACCAGGCCTGGCCA
G S A T G L A

TTCCGCCACCAGGCCTGGCCAT
S A T G L A

TTCCGCCACCAGGCCTGGCCAT
S A T G L A

TCCGCCACCAGGCCTGGCCATG
S A T G L A M

CCGCCACCAGGCCTGGCCATGA
S A T G L A M

CGCCACCAGGCCTGGCCATGAT
A T G L A M

GCCACCAGGCCTGGCCATGATC
A T G L A M I

CCACCAGGCCTGGCCATGATCA
T G L A M I

CACCCAGGCCTGGCCATGATCAC
T G L A M I

ACCCAGGCCTGGCCATGATCACA
T G L A M I T

Fig. 20 (cont'd 4)

T3

GGTCTGGTCAACCAAACAGAC
G L V N Q T D

GTCTGGTCAACCAAACAGACA
G L V N Q T D

TCTGGTCAACCAAACAGACAA
L V N Q T D

CTGGTCAACCAAACAGACAAG
L V N Q T D K

TGGTCAACCAAACAGACAAGG
L V N Q T D K

GGTCAACCAAACAGACAAGGA
V N Q T D K

GTCAACCAAACAGACAAGGAG
V N Q T D K E

TCAACCAAACAGACAAGGAGA
V N Q T D K E

CAACCAAACAGACAAGGAGAA
N Q T D K E

.CCAAACAGACAAGGAGAAA
N Q T D K E K

ACCAAACAGACAAGGAGAAAAG
N Q T D K E K

CCAAACAGACAAGGAGAAAAGG
Q T D K E K

CAAACAGACAAGGAGAAAAGGC
Q T D K E K G

AAACAGACAAGGAGAAAAGGCA
Q T D K E K G

AACAGACAAGGAGAAAAGGCAT
T D K E K G

ACAGACAAGGAGAAAAGGCATC
T D K E K G I

CAGACAAGGAGAAAAGGCATCT
T D K E K G I

.ACAAGGAGAAAAGGCATCTC
D K E K G I

GACAAGGAGAAAAGGCATCTCA
D K E K G I S

ACAAGGAGAAAAGGCATCTCAT
D K E K G I S

CAAGGAGAAAAGGCATCTCATC
K E K G I S

T3

102/124

TTCATGGATCCTCACTCTCCT
H G S S L S

TCATGGATCCTCACTCTCCTT
H G S S L S

CATGGATCCTCACTCTCCTTG
H G S S L S L

ATGGATCCTCACTCTCCTTGG
H G S S L S L

TGGATCCTCACTCTCCTTGGT
G S S L S L

GGATCCTCACTCTCCTTGGTT
G S S L S L V

GATCCTCACTCTCCTTGGTTT
G S S L S L V

ATCCTCACTCTCCTTGGTTTC
S S L S L V

TCCTCACTCTCCTTGGTTTCC
S S L S L V S

CCTCACTCTCCTTGGTTTCCA
S S L S L V S

CTCACTCTCCTTGGTTTCCAG
S L S L V S

TCCTCTCCTTGGTTTCCAGC
S L S L V S S

CACTCTCCTTGGTTTCCAGCA
S L S L V S S

ACTCTCCTTGGTTTCCAGCAC
L S L V S S

CTCTCCTTGGTTTCCAGCACA
L S L V S S T

TCTCCTTGGTTTCCAGCACAT
L S L V S S T

CTCCTTGGTTTCCAGCACATC
S L V S S T

TCCTTGGTTTCCAGCACATCG
S L V S S T S

CCTTGGTTTCCAGCACATCGT
S L V S S T S

CTTGGTTTCCAGCACATCGTC
L V S S T S

TTGGTTTCCAGCACATCGTCA
L V S S T S S

CTCCTTGGTTTCCAGCACATC
S L V S S T

TCCTTGGTTTCCAGCACATCG
S L V S S T S

CCTTGGTTTCCAGCACATCGT
S L V S S T S

CTTGGTTTCCAGCACATCGTC
L V S S T S

TTGGTTTCCAGCACATCGTCA
L V S S T S S

TGGTTTCCAGCACATCGTCAG
L V S S T S S

GGTTTCCAGCACATCGTCAGT
V S S T S S

GTTTCCAGCACATCGTCAGTT
V S S T S S V

TTTCCAGCACATCGTCAGTTT
V S S T S S V

TTCCAGCACATCGTCAGTTTA
S S T S S V

TCCAGCACATCGTCAGTTTAT
S S T S S V Y

CCAGCACATCGTCAGTTTATT
S S T S S V Y

CAGCACATCGTCAGTTTATTCT
S T S S V Y

AGCACATCGTCAGTTTATTCT
S T S S V Y S

GCACATCGTCAGTTTATTCTA
S T S S V Y S

CACATCGTCAGTTTATTCTAC
T S S V Y S

ACATCGTCAGTTTATTCTACA
T S S V Y S T

CATCGTCAGTTTATTCTACAC
T S S V Y S T

ATCGTCAGTTTATTCTACACC
S S V Y S T

TCGTCAGTTTATTCTACACCA
S S V Y S T P

CGTCAGTTTATTCTACACCAG
S S V Y S T P

T3

GGAAGAACTGGGTCAATGAGTTACGCAGCTCC
K N W V N E L R S S

Fig. 20 (cont'd 5)

T3

AAGAAGAAGAGGAAGAAGAACTGGGTCAATGAGTTACGCAGCTCCTTCAAG
K K K R K N W V N E L R S S F K

103/124

T3

TCTCTAATTTTCAATGCATGGATA

T3

AGGAGATGAAGCTGACAGATATCCGCTTAGAAGCTCT

T3

GATTCCAGACCACACGTCTTTCTTATCG

Fig. 20 (cont'd 6)

Note: The N-terminus of protein T2 was omitted in the alignment, since it has no significant homology to the T protein and the T3 protein.

T	---MDLSSEMNRHGKNPVSHKLEDQK-----KIYTDWANHYLAKSGHKRLIKDLQ	60
T3	NQPERLNSQVLQGLQEPAGEGLPLRKSGSVENGFDTQIYTDWANHYLAKSGHKRLIRDLO	
T2	-----	
	* * . . . * * . * . *****	
T	QDIADGVLLAEIIQIIANEKVEDINGCPRSQSQMIENVDVCLSFLAARGVNVQGLSAEEI	120
T3	QDVTGVLQAQIIQVVA NEKIEDINGCPKNRSQMIENIDACLNFLAAKGINIQGLSAEEI	
T2	-----	
	** . ***** . *** . ***** . ***** . ***** * ** ***** . * . *****	
T	RNGNLKAILGLFFSLSRKQQQ-HHQQQYYQS-----LVELQQRVT	180
T3	RNGNLKAILGLFFSLSRKQQQQQPQKQHLSSPLPPAVSQVAGAPSQCQAGTPQQQVPVT	
T2	-----	
	***** . * . * . *	
T	HASP-----PSEASQAKTQQDMQSRLPGP-SRVPAAGSSSKVQGASNLN--RRSQSFNSI	240
T3	PQAPCQPHQPAPHQQSKAQAEQSRPGPTARVSAAGSEAKTRGGSTTANNRRSQSFNNY	
T2	-----	
	. * . * . * . * . * . ***** . ** ***** . * . * . *****	
T	DKNK---PP-----	300
T3	DKSKPVTSPPPPPSSHEKEPLASSASSHPGMSDNAPASLESGSSSTPTNCSTSSAIPQPG	
T2	-----	
	** * **	
T	-----N-----	360
T3	AATKPWRSKSLSVKHSATVSMLSVKPPGPPEAPRPTPEAMKPAPNNQKSMLEKLKLFNSKG	
T2	-----	
	*	
T	-----	420
T3	GSKAGEGPGSRDTSCERLETLPSEEESELEAASRMLTTVGPASSSPKIALKGIAQRTFS	
T2	-----	
T	--YAN-----GNEK-----	480
T3	RALTNNKSSLKGNEKEKEKQREKDKEKSKDLAKRASVTERLDLKEEPKEDPSGAAPPEM	
T2	-----	
	. * *****	
T	-----	540
T3	PKKSSKIASFIPKGGKLN SAKKEPMAPSHSGIPKPGMKSM PGKSPSAPAPSKEGERSRSG	
T2	-----	
T	-----	600
T3	KLSSGLPQQKPQLDGRHSSSSSSSLASSEGGKPGGTTLNHSISSQTVSGSVGTTQTGTSNT	
T2	-----	

Fig. 21

T -----
T3 VSVQLPQPQQYNHPNTA APFLYRSQTDTEGNVTAESSSTGVSVEPFTKTGQPALE
T2 -----

T ---GEDPETRRMRTVKNIADLRQNL EETMSSLRGTQISHSTLETTFDSTVTTEVNGRTIP 720
T3 ELTGEDPEARRLRTVKNIADLRQNL EETMSSLRGTQVTHSTLETTFDTNVTTEMSGRSIL
T2 ----DPESQRKRTVQNVLDLRQNL EETMSSLRGSQVTHSSLEMTCYD--SDDANPRSVS
*****. * ****. * *****. * . * . * . *

T NLTSRPTPMTWRLGQACPRQLQAGDAPSLGAGYP-RSGTSRIFIHTDPSRFMYTTPLRRAAV 780
T3 SLTGRPTPLSWRLGQSSPRLQAGDAPSMGNGYPPRANASRFINTESGRYVYSAPLRRQLA
T2 SLSNRSYPLSWRYGQSSPRLQAGDAPSVGGSCRSEGTPAWYMHGERAHYSHTMPMR--SP
* . * * . * . * . * . * . * . * . * . * . *

T SRLGNMSQIDMSEKA-SSDL DMS-SEVDVGGYMSDGDILGKSLRTDDINSGYMTDGGLNL 840
T3 SRGSSVCHVDVSDKA-GDEMDLEGISMDAPGYMSDGDVLSKNIRTDDITSGYMTDGGLGL
T2 SKLSHISRLELVESLDSDEVDLK-----SGYMSDSDLMGKTMTEDDD---ITTG----
* * . * . * . * . * . *

T YTRSLNRIPD-TATSRDIIQRGVHDVTVDADSWDDSSSVSSGLSDTLDNISTDDLNTTSS 900
T3 YTRRLNRLPDGMVAVRET LQRNTSLGLGDADSWDDSSSVSSGISDTIDNLSTDDINTSSS
T2 -----WDESSSISSGLSDASDNLSSEEFNASSS
* . * . * . * . * . * . * . * . *

T VSSYSNITVPSRKN--TQLRTDSEKRSTTDET--WDSP--EELKKPE--EDFDSHG DAG- 960
T3 ISSYANTPASSRKN--LDVQTD AEKHSQVERNLSWSG---DDVKKSDGGSDSGIKMEPG-
T2 LNSLPSTPTASRRNSTIVLRTDSEKRSLAESGLSWFSESEEKAPKKLEYDSGLKMEPGT
. * * . * . * . * . * . * . *

T GKWKT VSSGLPEDPEK-AGQKASLSVSQTGSWRRGMSAQGGAPS---RQKAGTSALKTP- 1020
T3 SKWRRNP SDVSDSDKSTSGKKNPVISQTGSWRRGMTAQVGITMPRTKASAPAGALKTPG
T2 SKWRRERPESCDDSSKGGELKKPISLGHGPGSLKKGKTPPVAVTSP--ITHTAQSALKVAG
* . * . * . * . * . * . * . *

T -GKTDDAKASEKGKAPLKGSSQLRSPSDAGKSSGDEGKK--PPSGIGRSTATSSFGFKKP 1080
T3 TGKTDDAKVSEKGR LSPKASQVKRSPSDAGRSSGDESKKPLPSSSRTPTANANSFGFKKQ
T2 ---KPEGKATDKGKLAVKNTGLQRSSSDAGRDR L SDAKK--PPSGIARPSTSGSFGYKKP
. * * . * . * . * . * . * . *

T SG-VGSSAMITSSGATITSGSATLGKIPKSA AIGGKSNAGRKTS LDGSONQDDVVLHVSS 1140
T3 SGSATGLAMITASGVTVTSRSATLGKIPKSSALVSRS-AGRKSSMDGAQNQDDGYLALSS
T2 PP-ATGTATVMQTG-----GSATLSKI QKSSGIPVKPVNGRKTS LDVSN SAE PGFLAPGA
* . * * . * . * . * . * . *

T KTTLQYRSLPRPSKSSTSGIPGR-GGHRSTSSID-SNVSSKSAGATT SKLREPTKIGSG 1200
T3 RTNLQYRSLPRPSKSNSR--NG--AGNRSTSSID-SNISSKSAGLPV PKLREPSKTALG
T2 RSNIQYRSLPRPAKSSSMSVTGGRGGPRPVSSSIDP SLLSTKQGGLTPSRLKEPTKVASG
* . * . * . * . * . * . * . *

T RSSPVTVNQTDKEKEKVA VSDSES VSLSG-SPKSSPTSASACG-AQGLRQPGSKYPDIAS 1260
T3 SSLPGLVNQTDKEKG--ISSDNESVASCN-SVKVNPA AQPVSSPAQTS LQPGAKYPDVAS
T2 RTTPAPVNQTDREKE--KAKAKAVALDSDNISLSIGSPESTPKNQASHPTATKLAELP
. * * . * . * . * . *

T PTFRRLFGAKAGGKSASAPNTEGVKSSSVMPSPSTTLARQGSLESPSSGTGSMGSAGGLS 1320
T3 PTLRRLFGGKP-TKQVPIATAENMKN SVVISNPHATMTQQGNLDSPS-GSGVLS--S
T2 PTPLRAT-AKSFVKPPSLANLDKVN-SNSLDLPSSSDTTHASKVPDLHATSSAS-----
* . * . * . * . * . *

T GSSSPLFNKPSDLTTDV LSHSLASSPASVHSFTSGGLVWAANMSSSAGSKDTPSYQS 1380
 T3 GSSSPLYSKNVLDN-----QSPLASSPSSAHSAPSNSLTWGTNASSSSAVSKDGLGFQS
 T2 --GGPLPS-----CFTPSPAPILNINSASFSSQGLELMSGFVSPKETRMPYK
 ** * * * * *

 T MTSLHTSSESIDLPLS-----HHGSLSGLTTG-----THEVQSLLMRTGSRVSTLSES-- 1440
 T3 VSSLHTSCESIDISLSSGGVPSHNSSTGLIASS-----KDDSLTPFVRTNSVKTTLSESPL
 T2 LSGLHRSMESLQMPMSLPSAFPSSTPVPTPPAPPAAPTEETEELTWGSPRAGQLDS--
 .. ** * ** . . . * . . . *

 T -----MQLDRNTLPKKGLRYTPSSRQANQEEGKEWLRSHSTGGL 1500
 T3 SSPAASPKFCRSTLPRKQSDPHLDRNTLPKKGLRYTPTSQRLTQEDAKEWLRSHSAGGL
 T2 -----NQRDRNTLPKKGLRYQLQS----QEETKERRHSHTIGGL
 . ***** * ** * . . . *

 T QDTGNQSPVSPSAMSSSAAGKYHFSNLVSPTNLSQFNLPGPSMMRSNSIPAQDSSFDLY 1560
 T3 QDTAANSFPSSGSSVTSPSGTRFNFSQLASPTTVTQMSLSNPTMLRTHSLSNADGQYDPY
 T2 PESDDQSELPSPPALPMSLSAKGQLTNIVSPTAAT-----TPRITRSNSIPTHEAAFEY
 . . . * * ***** * . . . * . . . *

 T DDSQLCGSATSLEERPRAISHSGSFRDSMEEVHGSSLSLVSTSSLYSTAEKKAHSEQIH 1620
 T3 TDSRFRNSSMSLDEKSRTMSRSGSFRDGFEEVHGSSLSLVSTSSVYSTPEEKQCSE-IR
 T2 SGSQMG-STLSLAERPCKMIRSGSFRDPTDDVHGSVLSLASSASSTYSSAEERMQSEQIR
 * . * . * * . . . ***** . . . ***** * . . . * . . . *

 T KLRRELVASQEKVATLTSQLSANAHLVAAFEKSLGNMTGRLQSLTMTAEQKESELIELRE 1680
 T3 KLRRELDASQEKVSALTQTOLANAHLVAAFEQSLGNMTIRLQSLTMTAEQKDELNELRK
 T2 KLRRELESSQEKVATLTSQLSANANLVAAFEQSLVNMTSRLRHLAETAEEKDTELDDLRE
 ***** . . . ***** . . . ***** . . . ***** . . . *****

 T TIEMLKQAQNSAAQAAIQGALNGPDHPPK-----DLRIRRHSSSESVSSINSATSHSS 1740
 T3 TIELLKKQNAQAQAAINGVINTPELNCKGNGTAQSAADLRIRRHSSDSVSSINSATSHSS
 T2 TIDFLKKKNSAQAVIQGALNASETPPK-----ELRIKRONSSDSISLNSITSHSS
 ** . ** . * . * * * * . * . . . ***** . . . *****

 T IGSGNDADSKKKKKKNWL---RSSFKQAFGKKKSTKPPSSHSDIEELT--DSSLPASPKL 1800
 T3 VGSNIESDSKKKKRKNWVNELRSSFKQAFGKKKSPKSASSHSDIEEMT--DSSLPSSPKL
 T2 IGSSKDADAKKKKKKSWL---RSSFNKAFSIKKGPKSASSYSDIEEIATPDSSAPSSPKL
 . ** . . . ***** . * . . . ***** . * . . . *****

 T PHNAGDCGSASMKPSQSASAICECTEA-----EAEIILQLKSELRE 1860
 T3 PHNGSTGSTPLLRNHSNSLISECMDS-----EAETVMQLRNELRD
 T2 QHGSTETASPSIKSSTSSSVGTDVTEGPAHPAPHTRLFHANEEEEPEKKEVSELSELWE
 * * * * . . . * . . . *

 T KELKLTDIRLEALSSAHHLQDIREAMNRMQNEIEILKAENDRLKAETGNTAKPTRPPSES 1920
 T3 KEMKLTDIRLEALSSAHQLDQDREAMNRMQSEIEKLKAENDRLKSES-QGSGCSRAPSQV
 T2 KEMKLTDIRLEALNSAHQLDQDRETMHNMQLEVDLLEAENDRLKVAP--GPSSGSTPGQV
 ** . ***** * * . * * . * * . * * . * * . * * . * * . *

 T SSSTSSSSSRQSLGLSLNNLNITEAVSSDILLDDAGDATGHKDG-RSVKIIVSISKGYGR 1980
 T3 SISAS---PRQSMGLSQHSLNLTESTSLDMLLDDTGECSARKEGGRHVIVVSFQEMKW
 T2 PGSSALSSPRSLGLALTHSFGPSLADTDLSPMDGISTCGPKEE-VTLRVVVRMPPQHII
 * . . * . * . * . . * . . . * . . . *

```
T      MLMKLQEAANYSSSTQSCDSES--TSHHEDILDSSLESTL
T3     MLMRLQEAANYSSPQSYDSDSNSNSHHDDILDSSLESTL
T2     MLLKLQEAANYIE--SPDRET-----ILDPNLQATL
      **..*****      * * ..      ***  *..**
```

Fig. 21 (cont'd 3)

10 20 30 40 50 60

T-Protein
POM121

QDDVVLHVSSKTTTLQYRSLPRPSKSSTSGIPGRGGHRSSTSSIDSNVSSKSAGATTSKLR
QRKRKIQLLPSRRGDQLTLPPP-----P--ELG--YSITAEDLMERR---AS---LQ
* ** *

T-Protein
POM121 EPTKIGSGRSSPVTVNQTDKEKEKVAVSDSESVLSGSPKSSPTSASACGAQGLRQPGSK
WFKVLEDKTDASTPATDTSP---ATSPPTLTL---P---TVGPAASPASLPAPSS-
.. . . . * * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 YPDIASPTFRRLFGAKAGGKSASAPNTEGVKSSSVMPSPSTTLARQGSLESPSSGTGSMG
-----NPLLESLLKMQESPAPSSSEPPE--AATVAAPSPPKTPSLLAPLVSP-----
* * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 SAGGLSGSSSPLENKPSDLTDDVISLSHSLASSPASVHSFTSGGLVWAANMSSSSSAGSKD
----LTG---PLASTSSDSKPTTTFGLGLASASSATPLTDTKAPGVSAQLCVSTPAATAP
* * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 TPSYQSMSTSLHTSSESIDLPLSHHGLSGLTGTGTHEVQSLLMRTGSRVSTLSESMQLDRN
SP-----TPASTLFGMLSPASSSSSLATPGPACASPMFKPIFPATPKSE----SDN
* * * . . . * * . . . * * . . . *

T-Protein
POM121 TLPKKGLRYTPSSRQANQEEGKEWLRSHTSGGLQDTGNQSPLVSPSAMSSSAAGKYHFEN
PLP-----TSSSAATTPASTALPTTATATAHTFKPIFESVEPFAAMP-----
* * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 LVSPTNLSQFNLPGPSMMRSNSIPAQDSSFDLYDDSQLCGSATSLEERPRAISHSGSFRD
LSPPFSLKQTTAPATTAATSAPLLTG-----L-----GTATST-----VATGTTAS
* * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 SMEEVHGSSSLVVSSTSSLYSTAEKHAHQEQIHLRRELVASQEKVATLTSQLSANAHLV
ASKPVFGFGVTTAASTASTIAS-----TSQSILFGGAPPVTASSAPALASIFQFGKPLA
* * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 AAFEKSLGNMTGRLQSLTMTAEQKESELIELRETIEMLKQNSAAQAAIQGALNGPDHPP
PAASVAGTSFSSQSLASSAQTAAASNSS--GGFSGFGGTLTTSAPATTSQPTLTFSNTVT
* * * . . . * * . . . * * . . . *

T-Protein
POM121 KDLRIRRHSSSE-SVSSINSATSHSSIGSGNDADSKKKKKKNWLRSSFQKQAFGKKKSTK-
PTFNIPFSASAKPALPTYPGANSQPTFG-ATDGATKP-----ALAPSFSSFTFGNSVAS
* * * . . . * * . . . * * . . . *

T-Protein
POM121 PPSSHSIDIEELTDSSLPASPKLPENAGDCGSASMKPSQSASAIPECTEAEAEIILQKSE
APSAAPAPAAFGGAAQPAFGGLKASASTFG---TPASTQPAFGSTTS-----VFSFGSA
* * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 LREKELKLTDIRLEALSSAHHLQIREAMNRMONEIBILKAENDRLKAETGNTAKPTRPP
TTS-----GFGAAAATTQTHSGS-----SSSLFGSSTPS-PF
* * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 SESSSSTSSSSSRQSLGLSLNNLNITEAVSSDILLDDAGDATGHKDGSRVKIIVSISKGY
TFGGSAAAPAGG---GFGLSATPGTGSTSGTFSFGSGQSGT---TGTTTSFGGSLSQNT
* * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121 GRAKDQKSQAYLIGSIGVSGKTKWDVLDGVIRRLFKEYVFRIDTSTSLGLSSDCIASYCI
LGAPSQSS--PPAFSVGSTPESKP-----VFGGTSTPTFGQSAPAG---V
* * * . . . * * . . . * * . . . * * . . . *

T-Protein
POM121

GD LIRSHNLEVP E L P C G Y L V G D N N I I T V N L K G V E E N S L D S F V F D T L I P K P I T Q R Y F N L L
G T T G S S L S F G A P S T P A Q G F V G ----- V G P F G S G A P S F S I G A G S K T P G A R Q R L Q A R
* * * * *

T-Protein
POM121

M E H H R I I L S G P S G T G K T Y L A N K L A E Y V I T K S G R K K T E D A I A T F N V D H K S S K E L Q Q Y L A N L
R Q H T R K K -----
* *

T-Protein
POM121

A E Q C S A D N N G V E L P V V I I L D N L H H V G S L S D I F N G F L N C K Y N K C P Y I I G T M N Q G V S S S P N L

T-Protein
POM121

E L H H N F R W V L C A N H T E P V K G F L G R Y L R R K L I E I E I E R N I R N N D L V K I I D W I P K T W H H L N S

T-Protein
POM121

F L E T H S S S D V T I G P R L F L P C P M D V E G S R V W F M D L W N Y S L V P Y I L E A V R E G L Q M Y G K R T P W

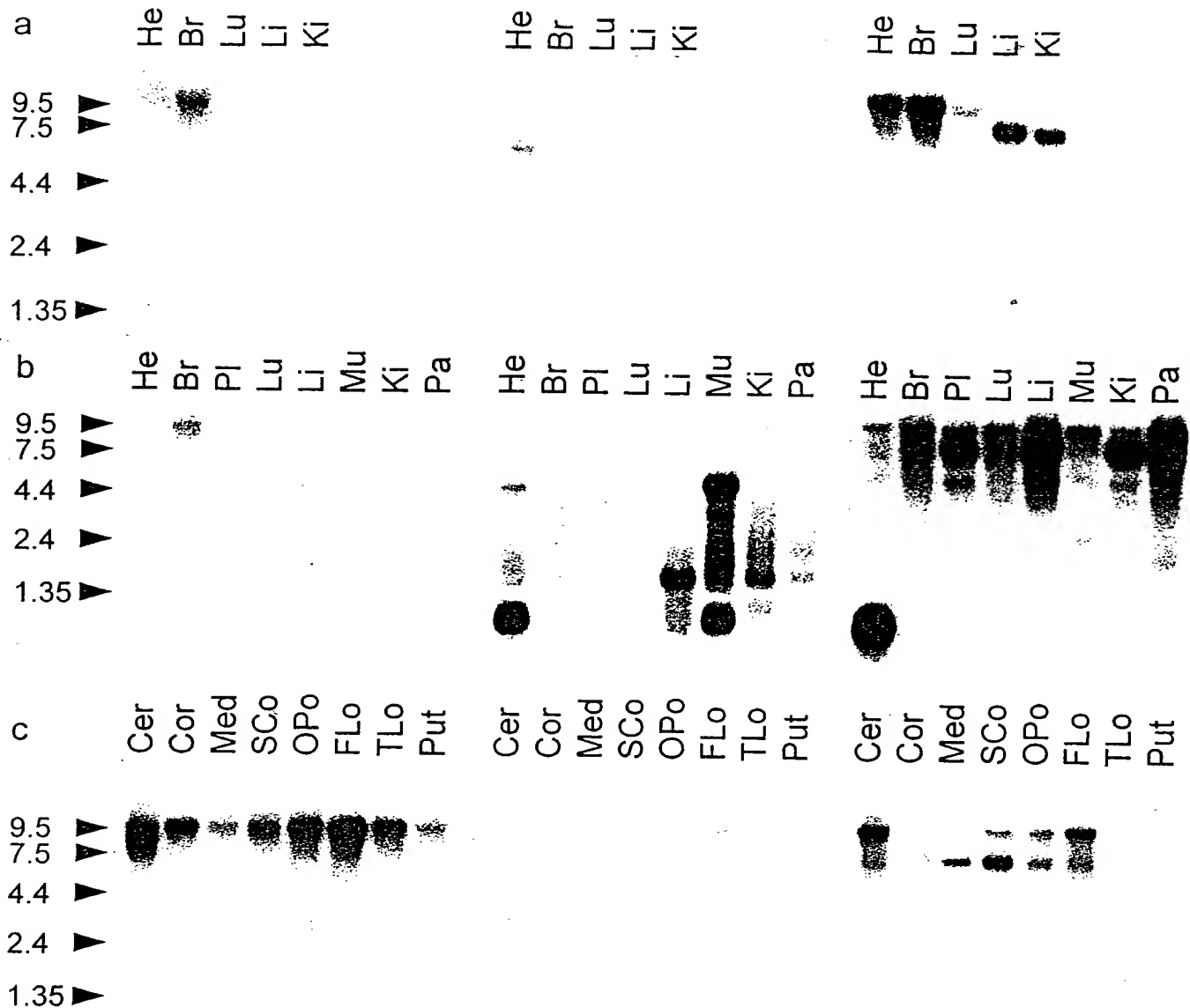
T-Protein
POM121

E D P S K W V L D T Y P W S S A T L P Q E S P A L L Q L R P E D V G Y E S C T S T K E A T T S K H I P Q T D T E G D P L

T-Protein
POM121

M N M L M K L Q E A A N Y S S T Q S C D S E S T S H H E D I L D S S L E S T L

Fig. 22 (cont'd 2)



Expression of the T gene family.

a fetal tissue: left: T gene; middle: T2 gene; right: T3 gene.

He = heart; Br = brain; Lu = lungs; Li = liver; Ki = kidney

b adult tissue: left: T gene; middle: T2 gene; right: T3 gene.

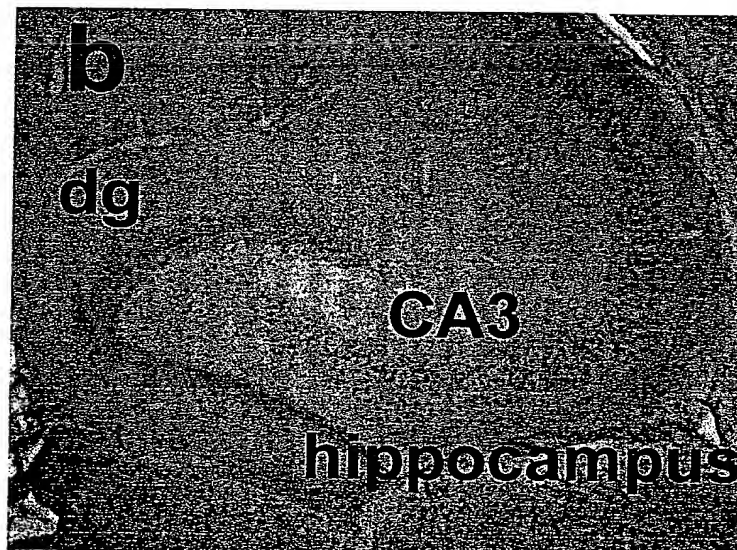
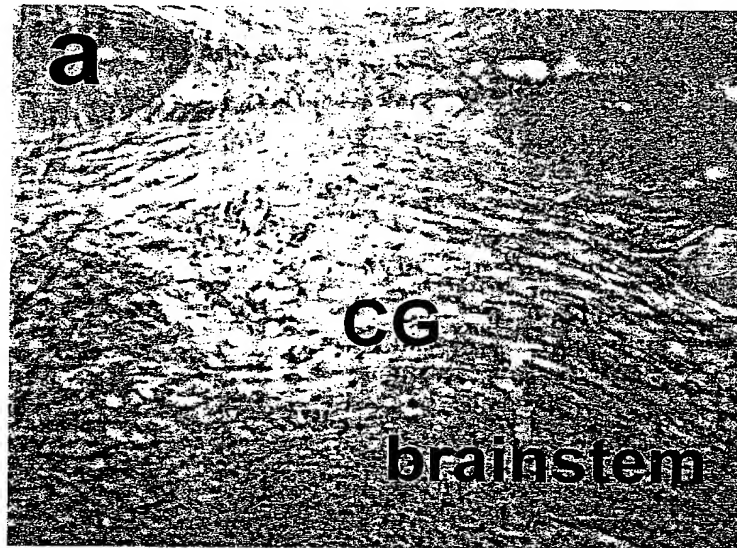
He = heart; Br = brain; Pl = placenta; Lu = lungs; Li = liver; Mu = skeletal muscle; Ki = kidney; Pa = pancreas

c adult brain regions: left: T gene; middle: T2 gene; right: T3 gene.

Cer = cerebellum; Cor = cerebral cortex; Med = medulla; Sco = spinal cord; Opo = occipital pole; Flo = frontal lobe; Tlo = temporal lobe; Put = putamen

Fig. 23

Fig. 24



113/124

Fig. 24

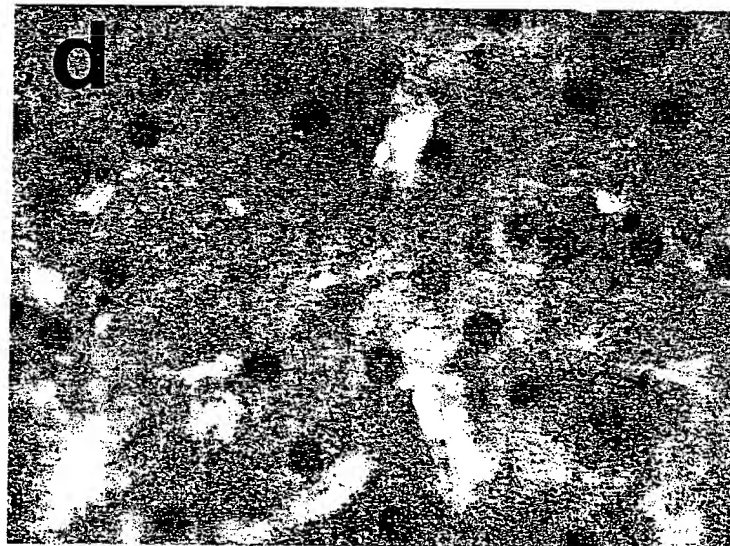
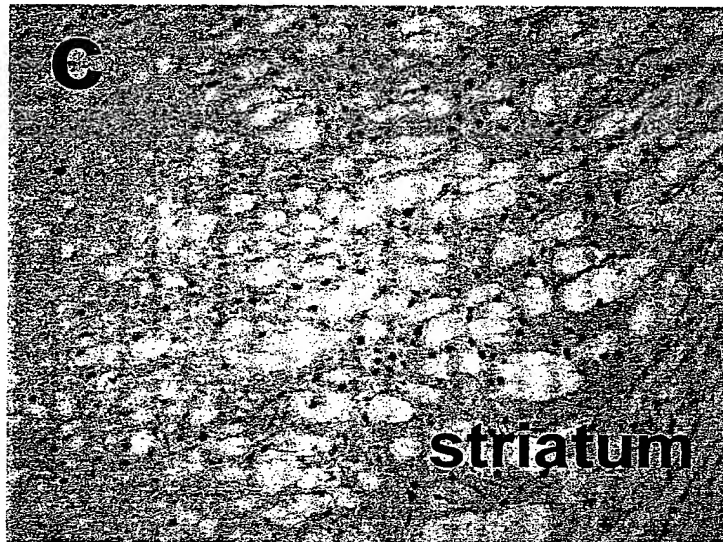


Fig. 24

114/124

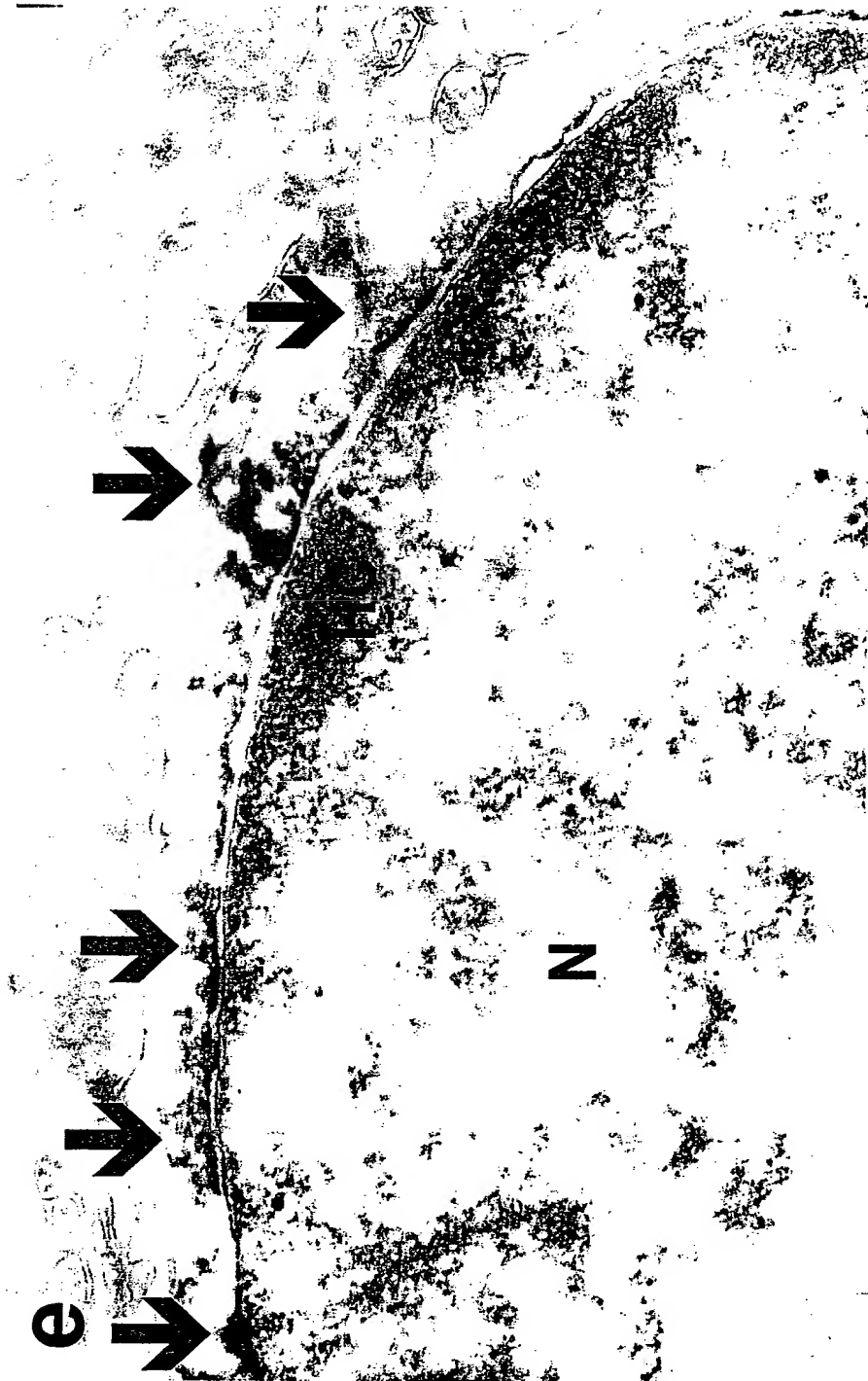
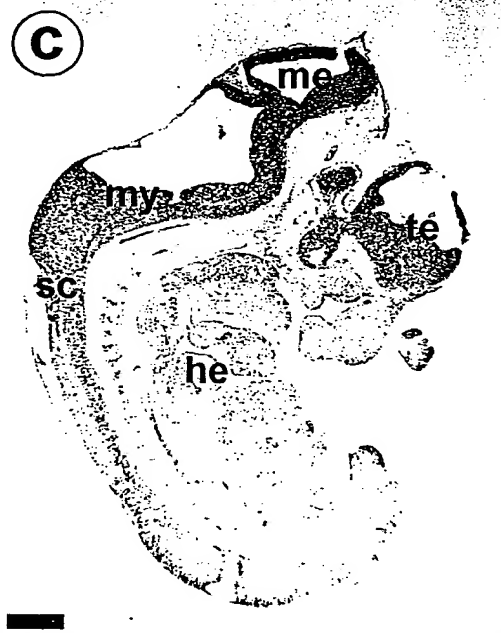
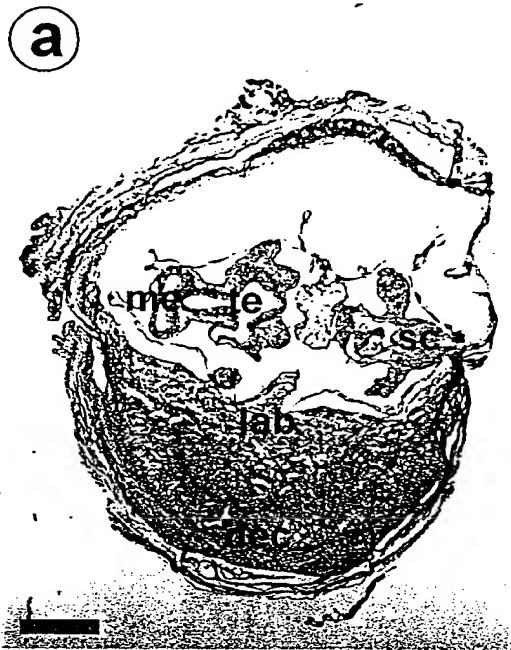


Figure legend of immunohisto and electron microscopy:

- a = brain stem. CG central grey = central grey of the brain stem
- b = hippocampus. dg = dental gyrus; CA3 cornu ammonis 3, both subregions of the hippocampus formation
- c = electronmicroscopic picture. N = nucleus, Hc heterochromatin

Fig. 25

115/124



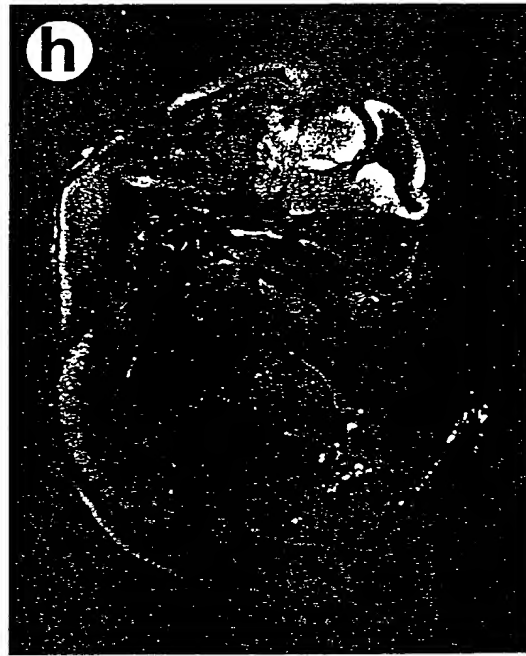
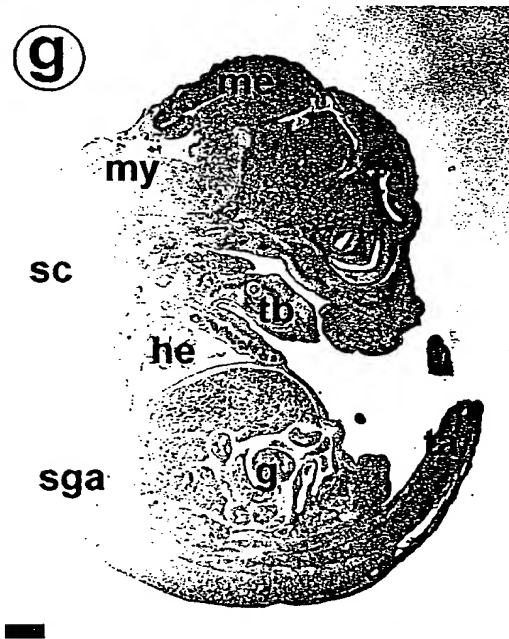
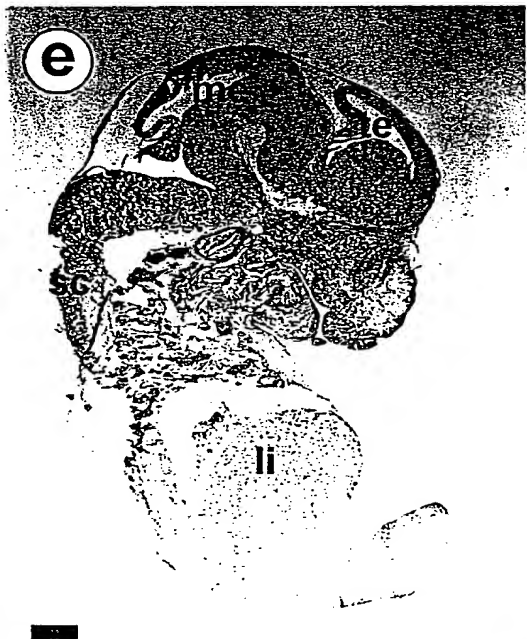


Fig. 26

117/124

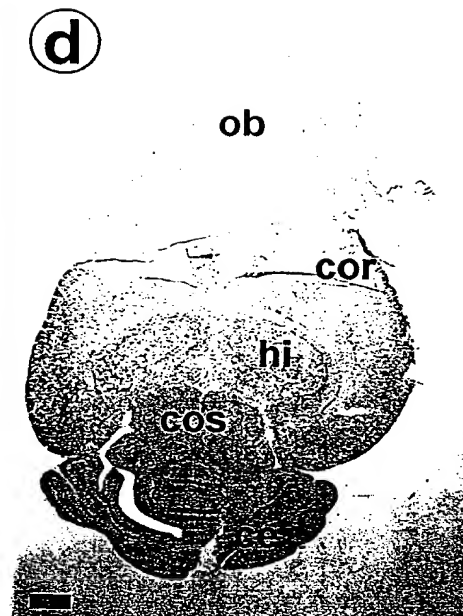
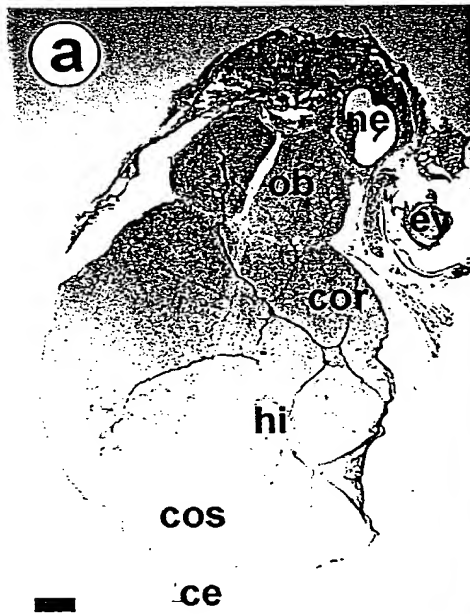


Fig. 26

118/124

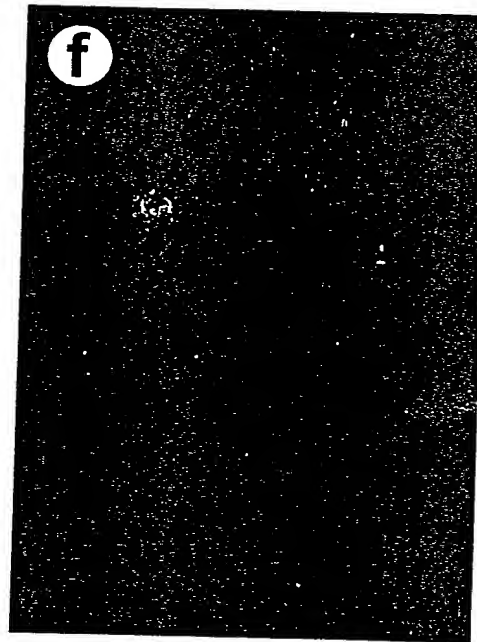


Fig. 27

119/124

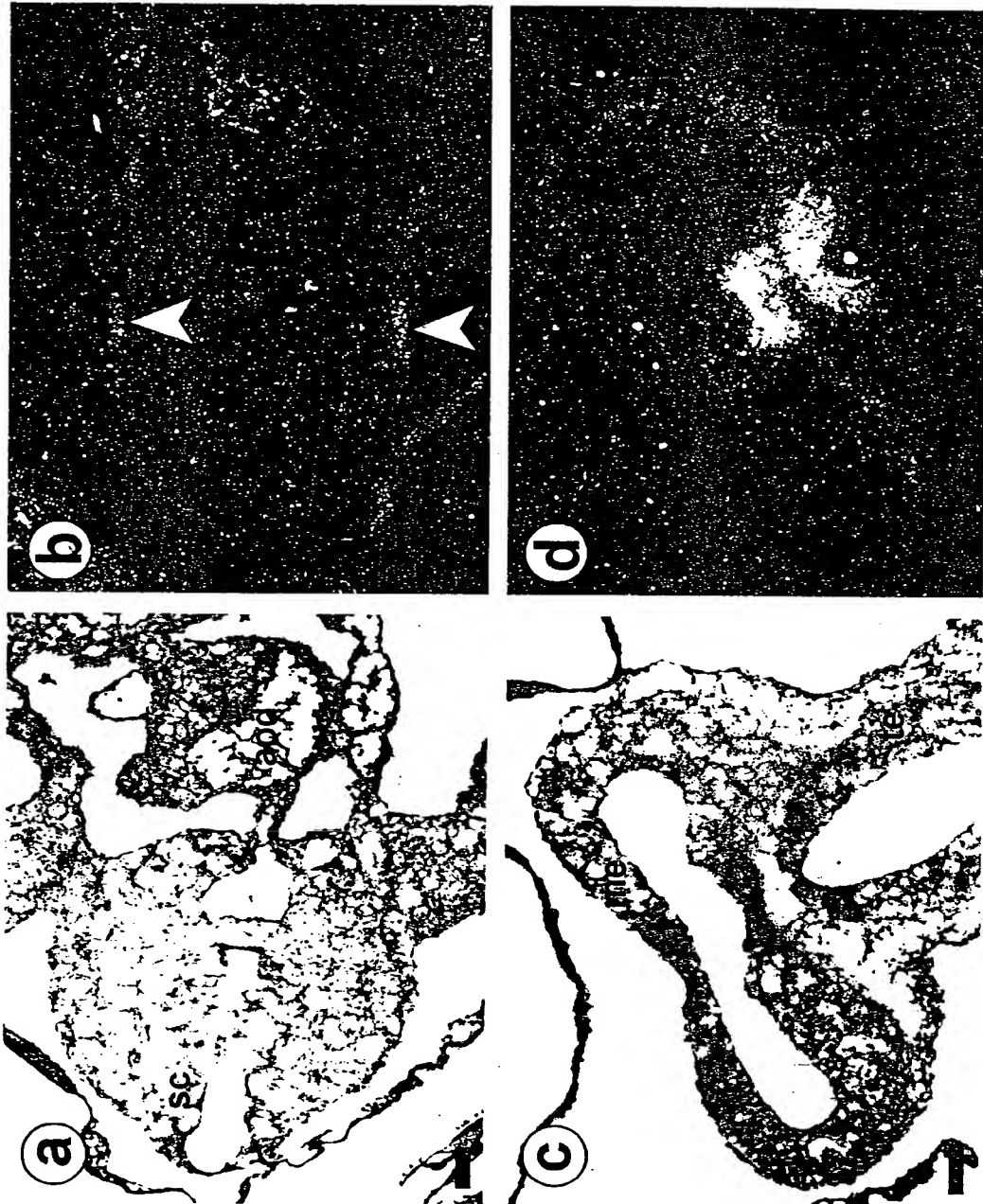


Fig. 27

120/124

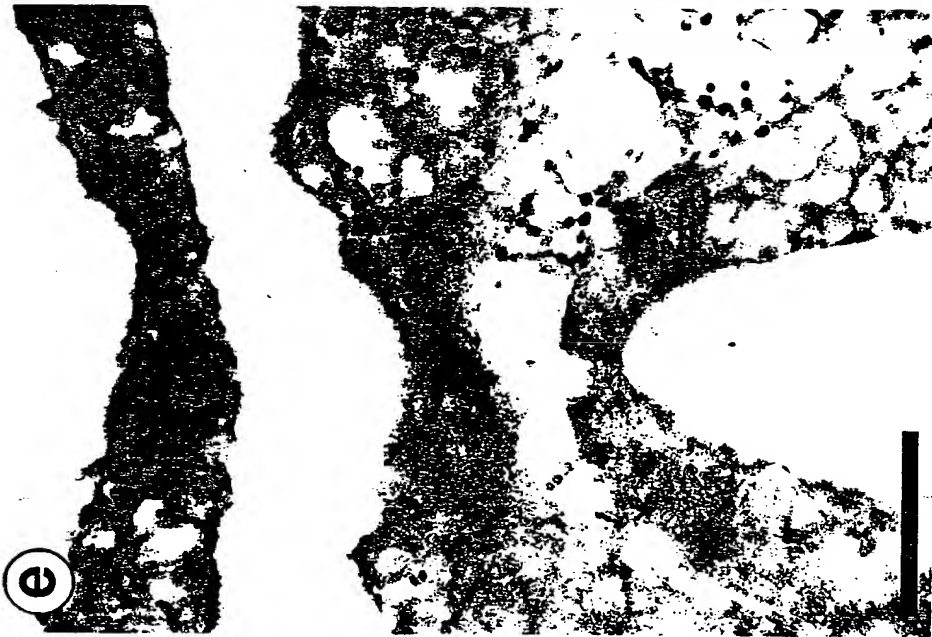
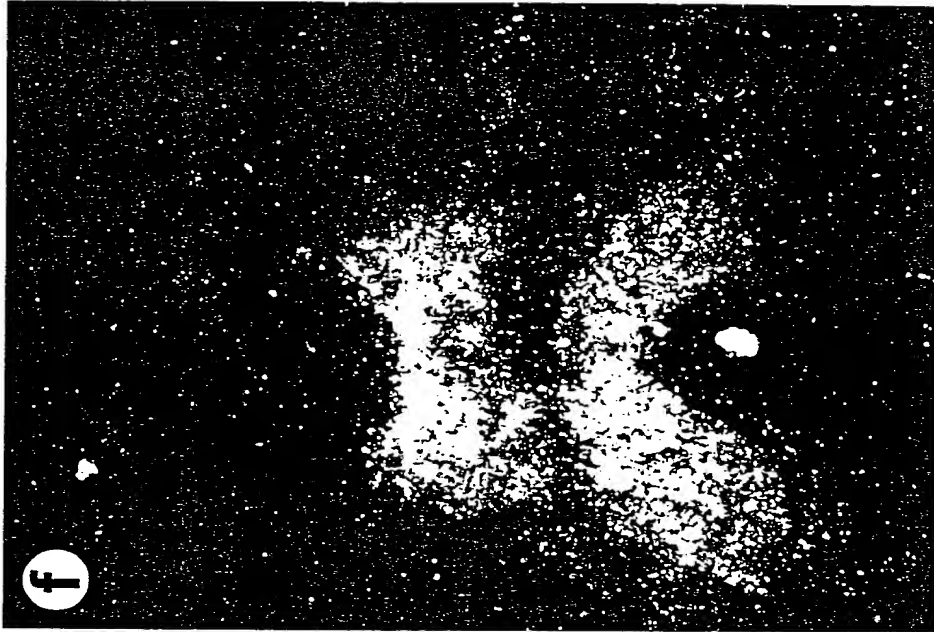


Fig. 28

121/124

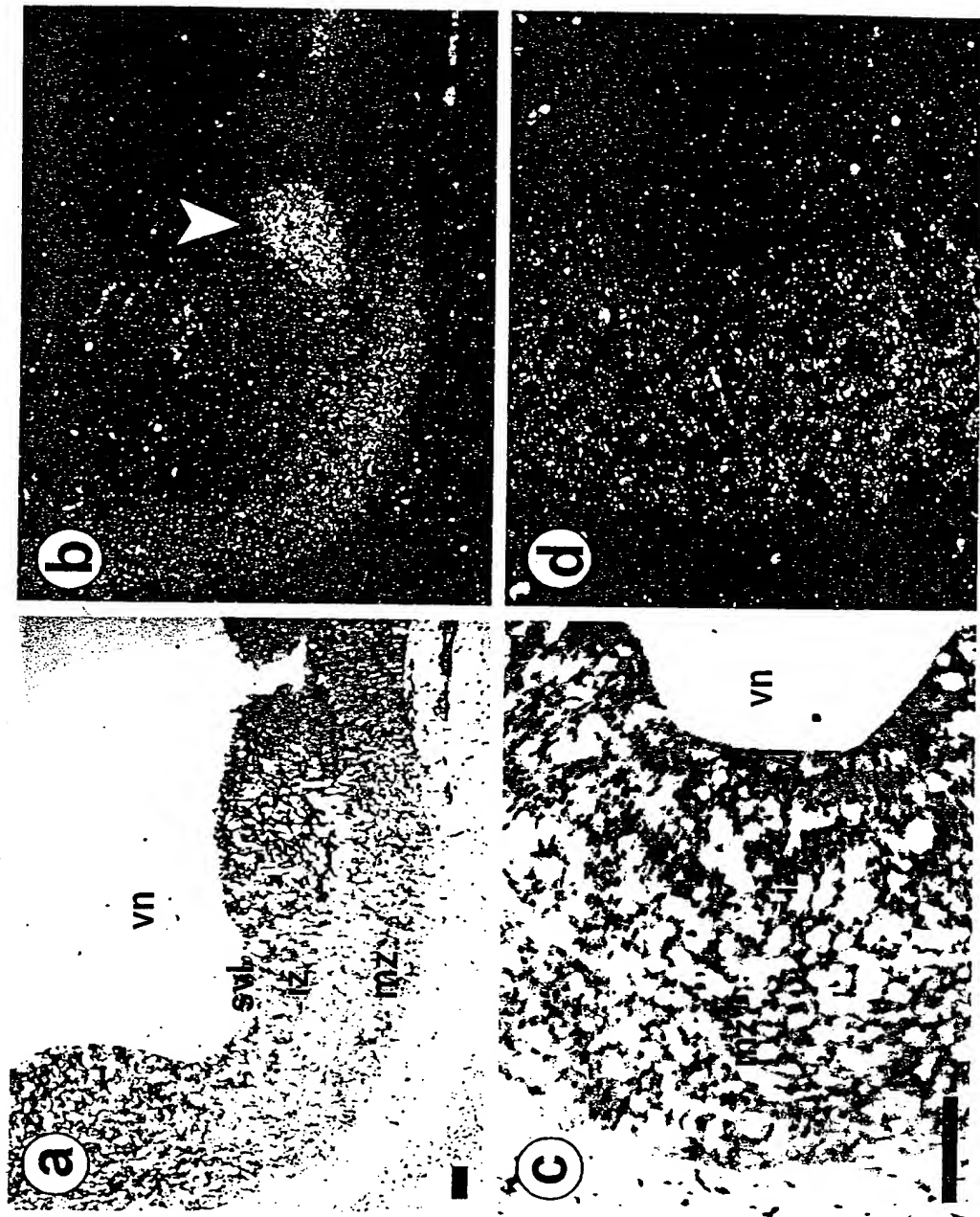


Fig. 28

122/124

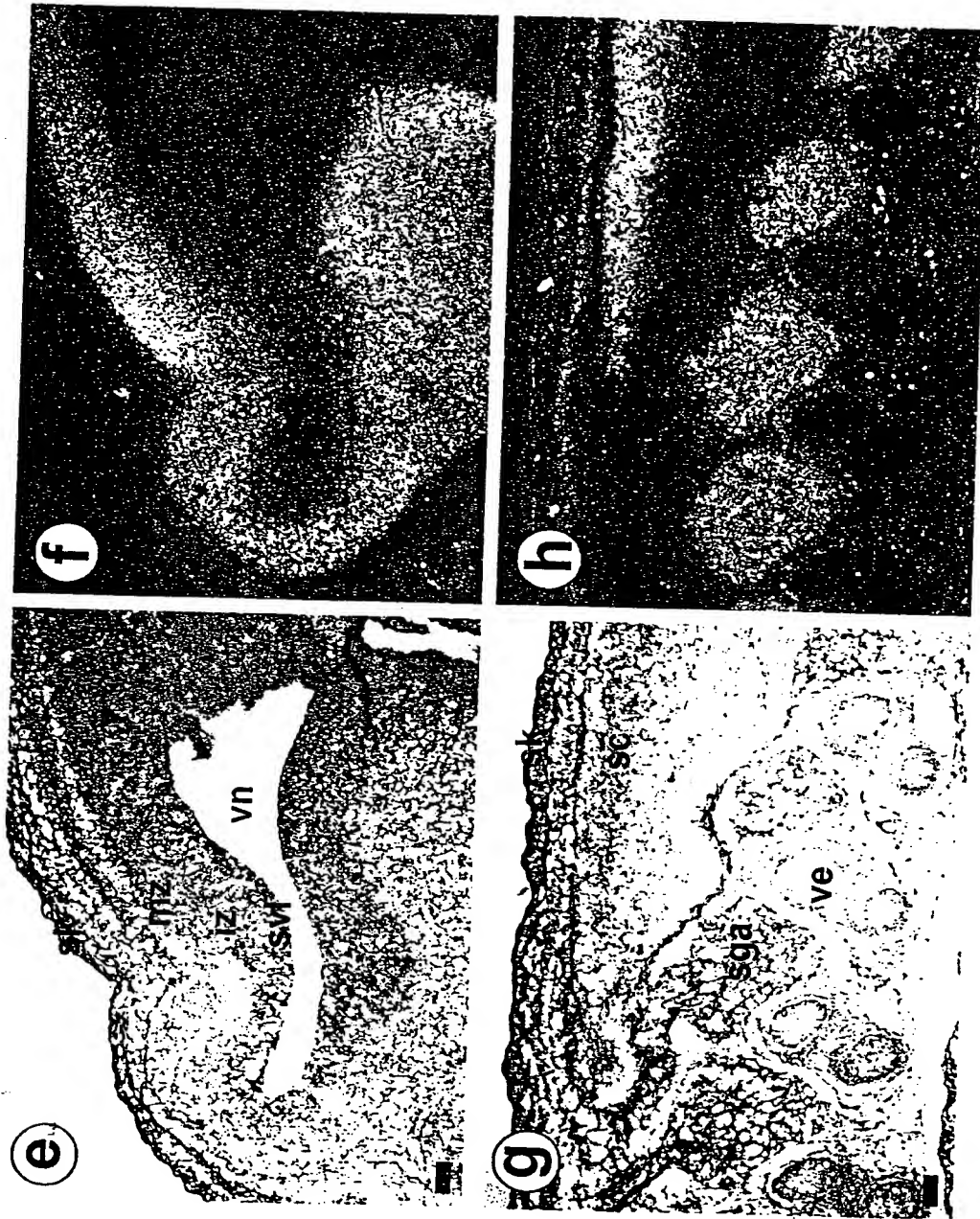
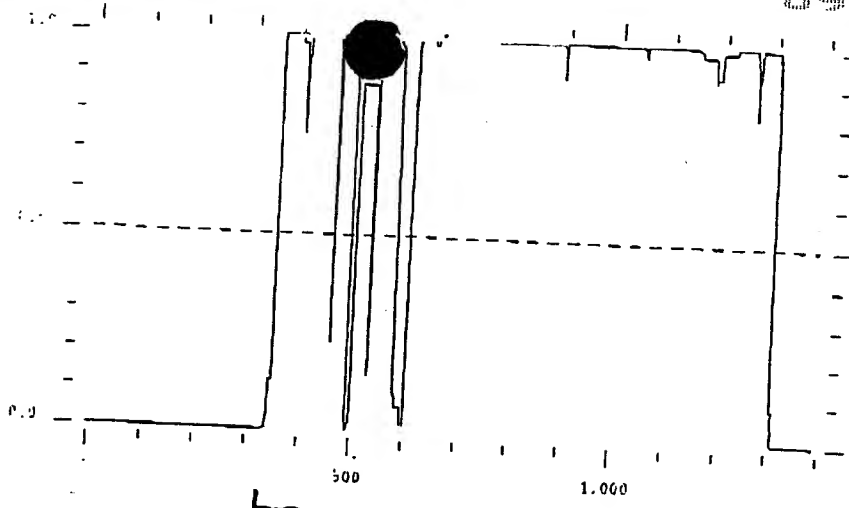


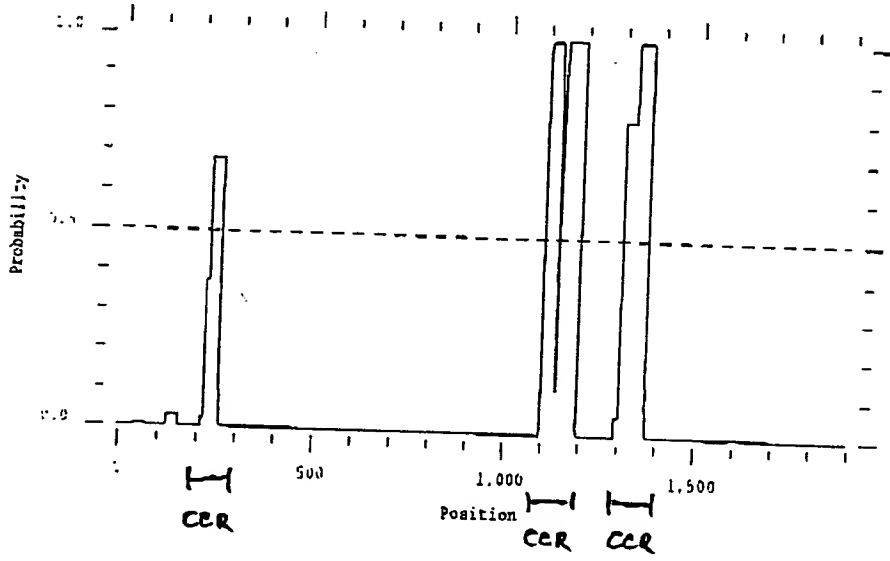
Fig. 29

CLIP-170



Coiled-Coil
Region (CCR)

T-Protein



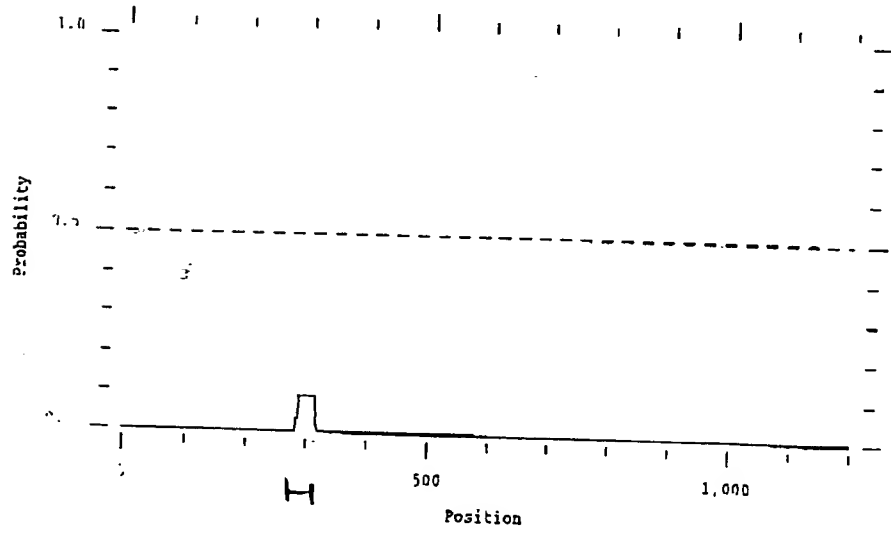
CCR

Position

CCR

CCR

POM 121



I

124/124

Fig. 30

